

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 23:02:20 ; Search time 775.333 Seconds
(without alignments)
1576.506 Million cell updates/sec

Title: US-09-765-111a-7

Perfect score: 42

Sequence: 1 agacctaccgcgtgtgca.....aatgaccatggttgacacag 42

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank

1: gb_ba:*

2: gb_hg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_ph:*

7: gb_pl:*

8: gb_pr:*

9: gb_ro:*

10: gb_sts:*

11: gb_sy:*

12: gb_un:*

13: gb_vt:*

14: gb_wi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_hg_hum:*

31: em_hg_inv:*

32: em_hg_other:*

33: em_hg_mus:*

34: em_hg_pln:*

35: em_hg_rod:*

36: em_hg_mam:*

37: em_hg_vrt:*

38: em_sy:*

39: em_hgo_hum:*

40: em_hgo_mus:*

41: em_hgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	30.2	71.9	1766	9	HSPARGAM	X90563 H.sapiens m
2	30.2	71.9	1808	9	HUMPPARGB	L40904 Homo sapien
3	30.2	71.9	1811	6	AX409565	AX409565 Sequence
4	30.2	71.9	1844	6	AR139020	AR139020 Sequence
5	25	59.5	170940	9	AC120042	AC120042 Homo sapi
6	24.6	58.6	33	6	AX151002	AX151002 Sequence
7	24	57.1	416	9	AB005521S1	AB005521 Homo sapi
8	24	57.1	135667	9	AC093174	AC093174 Homo sapi
9	24	57.1	166043	9	AC090947	AC090947 Homo sapi
10	24	57.1	185608	9	AC027126	AC027126 Homo sapi
11	23.4	55.7	209175	10	AL450321	AL450321 Nouse DNA
12	23.4	55.7	267260	2	AC099350	AC099350 Rattus no
13	23.2	55.2	191111	9	AC006080	AC006080 Homo sapi
14	23.2	55.2	194914	2	AC114359	AC114359 Rattus no
15	23	54.8	1518	9	HSU63415	U63415 Human perox
16	23	54.8	1608	6	AR121467	AR121467 Sequence
17	23	54.8	1608	9	HSU79012	U79012 Human ligan
18	23	54.8	1679	6	AR203332	AR203332 Sequence
19	23	54.8	1679	9	HUMPPARG	D83233 Homo sapien
20	23	54.8	60216	2	AC098574	AC098574 Oryza sat
21	23	54.8	128468	9	AL139090	AL139090 Human DNA
22	23	54.8	176072	2	AC120485	AC120485 Rattus no
23	22.8	54.3	1711	4	BTFFARG1	Y12419 B.taurus mR
24	22.8	54.3	161723	2	AC111044	AC111044 Mus muscu
25	22.8	54.3	176899	2	AC084108	AC084108 Mus muscu
26	22.8	54.3	192770	2	AC113481	AC113481 Mus muscu
27	22.8	54.3	213848	2	AC084401	AC084401 Mus muscu
28	22.6	53.6	98775	2	AC121723	AC121723 Rattus no
29	22.6	53.6	140596	2	RN75P15	AL603730 Rattus no
30	22.6	53.8	219471	2	AL772341	AL772341 Mus muscu
31	22.4	53.3	2028	9	AY048697	AY048697 Macaca fa
32	22.4	53.3	2063	9	AY048699	AY048699 Macaca fa
33	22.4	53.3	4065	12	AY071819	AY071819 Synthetic
34	22.4	53.3	39481	2	AC006176	AC006176 Homo sapi
35	22.4	53.3	42879	4	AC092249	AC092249 Canis fam
36	22.4	53.3	146276	4	AC090889	AC090889 Canis fam
37	22.4	53.3	152786	9	AL135778	AL135778 Human DNA
38	22.4	53.3	186204	2	AC102093	AC102093 Mus muscu
39	22.4	53.3	199785	9	AL355512	AL355512 Human DNA
40	22.4	53.3	233498	2	AC015882	AC015882 Homo sapi
41	22.4	53.3	237588	2	AC025581	AC025581 Mus muscu
42	22.2	52.9	163249	2	AC121362	AC121362 Oryza sat
43	22	52.4	346	6	AX340743	AX340743 Sequence
44	22	52.4	1372	9	HUMPA8A	L19606 Human paire
45	22	52.4	1380	4	CCPAX8A	X83591 C.canis Pax

ALIGNMENTS

RESULT 1	HSPARGAM	1766 bp	mRNA	linear	PRI 21-AUG-1997
LOCUS	H.sapiens mRNA for peroxisome proliferator activated receptor				
DEFINITION	H.sapiens mRNA for peroxisome proliferator activated receptor				
ACCESSION	X90563				
VERSION	X90563.1				
KEYWORDS	peroxisome proliferator-activated receptor gamma.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1766)				
AUTHORS	Lambe,K.G. and Tugwood,J.D.				
TITLE	A human peroxisome-proliferator-activated receptor-gamma is				

O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (02-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 5 (bases 1 to 170940)
 Biren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukigalter, B., Canarata, J., Chang, J., Chazato, B., Choepel, Y., Collamore, A., Cook, A., Cooke, P., DeArillano, K., Dewar, K., Diaz, O. S., Dodge, S., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Matthews, C., McCarthy, M., Weidman, J., Meneus, L., Minowa, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 2, 2002 this sequence version replaced gi:21327565.
 All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIGR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: L26752

Center clone name: 659_E_9

FEATURES

source

Location/Qualifiers

1..170940
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8"
 /clone="RP11-659E9"
 /clone_lib="RPC1-11 Human Male BAC"
 repeat_region 1..4156
 /rpt_family="L1PB1"
 repeat_region 4157..4186
 /rpt_family="L1PB1"
 repeat_region 4187..4887
 /rpt_family="L1PB1"
 repeat_region 4888..4940
 /rpt_family="L1PB1"
 repeat_region complement(4941..5012)
 /rpt_family="L1MCA"
 repeat_region 5019..6402
 /rpt_family="L1PB1"
 repeat_region 6443..6524
 /rpt_family="L1PB1"
 repeat_region 6525..6877
 /rpt_family="L1PB1"
 repeat_region 6878..6887
 /rpt_family="L1PB1"
 repeat_region complement(6888..7343)
 /rpt_family="L1M4"

repeat_region complement(7358..7728)
 /rpt_family="Charlie2"
 repeat_region complement(7734..7963)
 /rpt_family="L2"
 repeat_region complement(7993..8117)
 /rpt_family="Charlie2"
 repeat_region complement(8163..8442)
 /rpt_family="Charlie2"
 repeat_region 10205..10742
 /rpt_family="MLTIG1"
 repeat_region complement(10754..10847)
 /rpt_family="MIR"
 repeat_region complement(10908..11214)
 /rpt_family="AluSc"
 repeat_region 13348..13398
 /rpt_family="A-rich"
 repeat_region complement(13443..13496)
 /rpt_family="MER94"
 repeat_region 13662..13886
 /rpt_family="T)n"
 repeat_region 14096..14129
 /rpt_family="AT rich"
 repeat_region complement(14162..14479)
 /rpt_family="Aluub"
 repeat_region 14991..14933
 /rpt_family="(TTTTA)n"
 repeat_region complement(14536..14815)
 /rpt_family="AluSx"
 repeat_region 14899..14927
 /rpt_family="AT rich"
 repeat_region complement(15498..15819)
 /rpt_family="MER114"
 repeat_region 15848..15871
 /rpt_family="(CA)n"
 repeat_region complement(17438..17591)
 /rpt_family="MIR"
 repeat_region complement(19415..19502)
 /rpt_family="L2"
 repeat_region 19695..19856
 /rpt_family="L1MCI1"
 repeat_region complement(19857..20150)
 /rpt_family="AluY"
 repeat_region 20151..20758
 /rpt_family="L1MCI1"
 repeat_region 22151..22186
 /rpt_family="(TG)n"
 repeat_region 23453..23474
 /rpt_family="(CA)n"
 repeat_region 23601..23633
 /rpt_family="GA-rich"

Query Match 59.5%; Score 25; DB 9; Length 170940;
 Best Local Similarity 75.6%; Pred. No. 29;
 Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GACCTACCCCTGGTGGCAGAAATGACCATGGTTGACACAG 42
 Db 158464 GACCCACCCCAATGGTGTGAAATGACTATGGCAGACAG 158424

RESULT 6

AX151002
 LOCUS AX151002 33 bp DNA linear PAT 22-JUN-2001
 DEFINITION Sequence 1 from Patent WO0138325.
 ACCESSION AX151002
 VERSION AX151002.1 GI:14533229
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 REFERENCE 1 (bases 1 to 33)
 AUTHORS Momose, Y., Maekawa, T., Odaka, H. and Kimura, H.
 TITLE 5-membered n-heterocyclic compounds with hypoglycemic and

hypolipidemic activity
JOURNAL Patent: WO 0138225-A 1 31-MAY-2001;
Takeda Chemical Industries, Ltd. (JP)
FEATURES Location/Qualifiers
source
1. .33
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Primer for cloning ppar? gene"

BASE COUNT 10 a 6 c 11 g 6 t
ORIGIN

Query Match 58.6%; Score 24.6; DB 6; Length 33;
Best Local Similarity 87.1%; Pred. No. 54;
Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 12 GTGTGGCAGAAATGACCATGGTTGACACAG 42
|||||
Db 1 GTGGCTACCGAAATGACCATGGTTGACACAG 31
|||||

RESULT 7
AB005521S1
LOCUS Homo sapiens ppar gamma gene for peroxisome proliferator
DEFINITION activated-receptor gamma, exon 1.
ACCESSION AB005521
VERSION 1
KEYWORDS ppar gamma; ppar gamma common exon1; peroxisome proliferator
SEGMENT 1 of 6
SOURCE activated-receptor gamma.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Okazawa,H., Mori,H., Tamori,Y., Araki,S., Niki,T., Masugi,J.,
Kawanishi,M., Kubota,T., Sinoda,H. and Kasuga,M.
TITLE No coding mutations are detected in the peroxisome
proliferator-activated receptor- gene in Japanese patients with
lipotrophic diabetes
JOURNAL Diabetes (1997) In press
REFERENCE 2 (bases 1 to 416)
AUTHORS Okazawa,H.
TITLE Direct Submission
SUBMITTED (03-JUL-1997) Hideki Okazawa, Kobe University School of
Medicine, 2nd Department of Internal Medicine; 7-5-1 Kusumoki-cho
cho-ku, Kobe 650, Japan (E-mail:okazawa@med.kobe-u.ac.jp,
Tel:81-78-341-7451, Fax:81-78-382-2080)
FEATURES Location/Qualifiers
source
1. .416
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="placenta"
<115..340
/gene="ppar gamma"
/product="peroxisome proliferator activated-receptor
gamma"
/note="ppar gamma common exon1"
/number=1

BASE COUNT 113 a 105 c 67 g 131 t
ORIGIN

Query Match 57.1%; Score 24; DB 9; Length 416;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 CAGAAATGACCATGGTTGACACAG 42
|||||
Db 110 CAGAAATGACCATGGTTGACACAG 133
|||||

RESULT 8
AC093174
LOCUS Homo sapiens chromosome 3 clone RP11-167M22 map 3p, complete
DEFINITION sequence.
ACCESSION AC093174
VERSION 1
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 135667)
AUTHORS Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D., Guo,Z.,
He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,F.,
Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y.,
Li,W., Li,W., Li,Y., Luo,J., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L.,
Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
Wang,J., Wang,L., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X.,
Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X., Yu,B.,
Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
Yu,Q. and Yang,H.

TITLE Chromosome 3p genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 135667)
AUTHORS Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D., Guo,Z.,
He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,F.,
Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y.,
Li,W., Li,W., Li,Y., Luo,J., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L.,
Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
Wang,J., Wang,L., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X.,
Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X., Yu,B.,
Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
Yu,Q. and Yang,H.

TITLE Direct Submission
JOURNAL Submitted (13-AUG-2001) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
COMMENT -----Genome Center
Center:Beijing Center
Center code:Beijing
Website:http://hgsc.igtp.ac.cn
http://www.genomics.org.cn
Contact:hgsc@igtp.ac.cn
----- Project Information
Center project name:1% project
Center clone name: RP11-167M22
----- Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator: Et 53% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 586 bases at least Q40
Consensus quality: 919 bases at least Q30
Consensus quality: 1154 bases at least Q20
Insert size: 1198; sum-of-contigs
Quality coverage: 1.48x in Q20 bases;sum-of-contigs

FEATURES Location/Qualifiers
source
1. .135667
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="RP11-167M22"
BASE COUNT 41244 a 26375 c 26554 g 42494 t
ORIGIN

Query Match 57.1%; Score 24; DB 9; Length 135667;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 CAGAAATGACCATGGTTGACACAG 42

Db 109924 CAGAAATGACCATGGTTGACACAG 109947

RESULT 9
AC090947/c
LOCUS AC090947 166043 bp DNA linear PRI 20-MAR-2001
DEFINITION Homo sapiens chromosome 3 clone RP11-30G23 map 3p, complete sequence.

ACCESSION AC090947, AC016333
VERSION AC090947.1 GI:13384351
KEYWORDS HTG.
SOURCE Homo sapiens.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 166043)
AUTHORS Wu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, W., Li, Y., Luo, C., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.

Chromosome 3p genomic sequence

Unpublished

2 (bases 1 to 166043)
AUTHORS Wu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, W., Li, Y., Luo, C., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.

Direct Submission
Submitted (20-MAR-2001) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China

On Mar 20, 2001 this sequence version replaced gi:8072582.

-----Genome Center

Center:Beijing Center

Center code:Beijing

Website: <http://hgsc.igtp.ac.cn>

<http://www.genomics.org.cn>

Contact: hgsc@igtp.ac.cn

-----Project Information

Center project name: 1k project

Center clone name: RP11-30G23

-----Summary Statistics

Sequencing vector: pUC18; 100% of reads

Chemistry: Dye-terminator; ET 55% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 499 bases at least Q40

Consensus quality: 605 bases at least Q30

Consensus quality: 674 bases at least Q20

Insert size: 692; sum-of-contigs

Quality coverage: 2.80x in Q20 bases; sum-of-contigs

Location/Qualifiers

1. 166043

/organism="Homo sapiens"

/db_xref="taxon:9606"

FEATURES

source

/chromosome="3"

/map="3p"

/clone="RP11-30G23"

BASE COUNT 48570 a 35113 c 34446 g 47914 t

ORIGIN

Query Match 57.1%; Score 24; DB 9; Length 166043;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 CAGAAATGACCATGGTTGACACAG 42

Db 117076 CAGAAATGACCATGGTTGACACAG 117053

RESULT 10

AC027126/c

LOCUS AC027126 185608 bp DNA linear PRI 06-MAR-2001

DEFINITION Homo sapiens chromosome 3 clone RP11-586C12 map 3p, complete sequence.

ACCESSION AC027126

VERSION AC027126.4 GI:13236635

KEYWORDS HTG.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 185608)

AUTHORS Wu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, W., Li, Y., Luo, C., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.

Chromosome 3p genomic sequence

Unpublished

2 (bases 1 to 185608)

AUTHORS Tao, R., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B., Bao, W., Sun, Y., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y., Niu, Y., Qi, X., Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y., Li, G., Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D., Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J. and Yang, H.

Direct Submission

Submitted (28-MAR-2000) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China

3 (bases 1 to 185608)

AUTHORS Wu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, W., Li, Y., Luo, C., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.

Chromosome 3p genomic sequence

Unpublished

2 (bases 1 to 185608)

AUTHORS Tao, R., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B., Bao, W., Sun, Y., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y., Niu, Y., Qi, X., Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y., Li, G., Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D., Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J. and Yang, H.

Direct Submission

Submitted (28-MAR-2000) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China

3 (bases 1 to 185608)

AUTHORS Wu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, W., Li, Y., Luo, C., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.

Chromosome 3p genomic sequence

Unpublished

2 (bases 1 to 185608)

AUTHORS Tao, R., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B., Bao, W., Sun, Y., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y., Niu, Y., Qi, X., Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y., Li, G., Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D., Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J. and Yang, H.

Direct Submission

Submitted (28-MAR-2000) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China

3 (bases 1 to 185608)

AUTHORS Wu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, W., Li, Y., Luo, C., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.

Chromosome 3p genomic sequence

Unpublished

2 (bases 1 to 185608)

AUTHORS Tao, R., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B., Bao, W., Sun, Y., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y., Niu, Y., Qi, X., Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y., Li, G., Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D., Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J. and Yang, H.

Direct Submission

Submitted (28-MAR-2000) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China

3 (bases 1 to 185608)

AUTHORS Wu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, W., Li, Y., Luo, C., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.

Chromosome 3p genomic sequence

Unpublished

2 (bases 1 to 185608)

AUTHORS Tao, R., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B., Bao, W., Sun, Y., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y., Niu, Y., Qi, X., Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y., Li, G., Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D., Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J. and Yang, H.

Direct Submission

Submitted (28-MAR-2000) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China

3 (bases 1 to 185608)

AUTHORS Wu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, W., Li, Y., Luo, C., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.

Chromosome 3p genomic sequence

Unpublished

2 (bases 1 to 185608)

AUTHORS Tao, R., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B., Bao, W., Sun, Y., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y., Niu, Y., Qi, X., Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y., Li, G., Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D., Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J. and Yang, H.

Direct Submission

Submitted (28-MAR-2000) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing 100101, P.R.China

3 (bases 1 to 185608)

AUTHORS Wu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F., Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y., Li, W., Li, W., Li, Y., Luo, C., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L., Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J., Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X., Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N., Yu, J. and Yang, H.

Chromosome 3p genomic sequence

Unpublished

2 (bases 1 to 185608)

AUTHORS Tao, R., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B., Bao, W., Sun, Y., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y., Niu, Y., Qi, X., Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y., Li, G., Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D., Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J. and Yang, H.

Direct Submission

<http://www.genomics.org.cn>

Contact: hgc@igtp.ac.cn

----- Project Information

Center project name: 111 project

Center clone name: RP11-586C12

----- Summary Statistics

Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator; 55% of reads
Assembly program: Phrap; version 0.930329

Consensus quality: 187858 bases at least Q40

Consensus quality: 189056 bases at least Q30

Consensus quality: 189280 bases at least Q20

Insert size: 185608; sum-of-contigs

Quality coverage: 10.96x in Q20 bases; sum-of-contigs

FE ES source

----- Location/Qualifiers

1. 185608

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="3"

/map="3p"

/clone="RP11-586C12"

BASE COUNT 53641 a 40367 c 39222 g 52378 t

ORIGIN

Query Match

Best Local Similarity 57.1%; Score 24; DB 9; Length 185608;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 CAGAAATGACCATGGTTGACACAG 42

Db 161293 CAGAAATGACCATGGTTGACACAG 161270

RESULT 11

AL450321/c

LOCUS

DEFINITION Mouse DNA sequence from clone Rp23-232M10 on chromosome 13,

complete sequence.

ACCESSION AL450321

VERSION AL450321.15 GI:14456245

KEYWORDS HTG.

SOURCE house mouse.

ANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 209175)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (29-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Jun 15, 2001 this sequence version replaced gi:14330010.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em: EMBL; Sw:

SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP

database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-232M10 is

from the RP1-23 Mouse PAC Library

constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBac3.6

----- Genome Center

Center: UK Medical Research Council

Center code: UK-MRC

Web site: <http://mrcseq.har.mrc.ac.uk>

Contact: mouse@har.mrc.ac.uk

FEATURES

source

----- Location/Qualifiers

1. 209175

/organism="Mus musculus"

/db_xref="taxon:10090"

/chromosome="13"

/clone="RP23-232M10"

/clone_lib="RP1-23"

BASE COUNT 56180 a 44574 c 46218 g 62203 t

ORIGIN

Query Match

Best Local Similarity 55.7%; Score 23.4; DB 10; Length 209175;

Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GACCTACCCGCTGGTGGCAGAAATGACCATGGTTGACACAG 42

Db 114437 GACACACCTTGTGGTGGTTTAATAAGATGGTTTACACAG 114397

RESULT 12

AC099350

LOCUS

DEFINITION Rattus norvegicus clone CH230-6013, *** SEQUENCING IN PROGRESS ***

ACCESSION AC099350

VERSION AC099350.4 GI:21717844

KEYWORDS HTG; HTGS PHASE1.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 267260)

AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,

Alsbrooks, S.L., Amaratunga, H.C., Are, J.R., Ayele, M., Banks, T.,

Barbaris, J., Beaton, J., Binage, K., Blankenburg, K., Bonnin, D.,

Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,

Buhay, C., Burck, P., Burkett, C., Burrell, K.L., Byrd, N.C.,

Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,

Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,

Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,

Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,

Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,

Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,

Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,

Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,

Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,

Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,

Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,

Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,

Homes, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,

Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,

Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,

Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,

Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Louleghed, H.,

Lozato, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,

Maheeshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, B.,

Massey, E., Mahoney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,

Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,

Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,

Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,

Oren, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,

Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,

Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,


```

* 154210 163071: contig of 8862 bp in length
* 163072 163171: gap of unknown length
* 163172 171897: contig of 8726 bp in length
* 171897 171997: gap of unknown length
* 171998 182403: contig of 10406 bp in length
* 182403 182503: gap of unknown length
* 182404 193522: contig of 11019 bp in length
* 182504 193522: gap of unknown length
* 193523 210701: contig of 17079 bp in length
* 193623 210701: gap of unknown length
* 210702 236951: contig of 26150 bp in length
* 210802 236951: contig of 26150 bp in length

Query Match      55.7%; Score 23.4; DB 2; Length 267260;
Best Local Similarity 73.2%; Pred. No. 1.2e+02;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Q: 1 AGACCTACCCCGTGGTGAGAAATGACCATGGTTGACACA 41
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 26039 AGACACACCCCTGGTGGCGCACCGACATGGTGTGACACA 26079

RESULT 13
AC006080/c
LOCUS          AC006080      191111 bp      DNA      linear      PRI 23-DEC-1998
DEFINITION    Homo sapiens chromosome 17, clone hRPK.394_K10, complete sequence.
ACCESSION     AC006080
VERSION       AC006080.1  GI:4056515
KEYWORDS      HTG.
SOURCE        Homo sapiens.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 191111)
AUTHORS       Birren,B., Linton,L., Nussbaum,C., Lander,E., Allen,N., Anderson,M.,
               Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Boutwell,C.,
               Brown,A., Castle,A., Cerny,J., Collangelo,M., Collins,S.,
               Collymore,A., Cooke,P., Corliss,D., Depayre,E., Devon,K., Dewar,K.,
               Donegan,L., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R.,
               Gage,D., Gardyna,S., Grainger,K., Grant,G., Hagos,B., Heaford,A.,
               Herena,L., Horton,L., Howland,J.C., Jacotot,L., Jones,C., Kann,L.,
               Karatas,A., Lehoczyk,J., Macdonald,P., Marquis,N., McEwan,P.,
               McGurk,A., McKernan,K., Meldrim,J., Mollia,M., Morris,W., Morrow,J.,
               Mychaleckyj,J., Nahf,R., Naylor,J., Niloff,M., O'Connor,T., Roy,A.,
               O'Donnell,P., Pavlin,B., Peterson,K., Riley,R., Roberts,D., Roy,A.,
               Severy,P., Stange-Thomann,N., Scilwell,J., Stojanovic,N., Stone,C.,
               Subramanian,A., Tesfaye,S., Tichovolsky,N., Torruella-Miller,I.,
               Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,Y., Wyman,D.,
               Ye,W.J., Zhao,J. and Zody,M.
               Direct Submission
               Submitted (01-DEC-1998) Whitehead Institute/MIT Center for Genome
               Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE     3 (bases 1 to 191111)
AUTHORS       Birren,B., Linton,L., Nussbaum,C., Lander,E., Allen,N., Anderson,M.,
               Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
               Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
               Cooke,P., DeAtellano,K., Depayre,E., Devon,K., Dewar,K.,
               Donegan,L., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R.,
               Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A.,
               Herena,L., Horton,L., Howland,J.C., Jones,C., Kann,L., Karatas,A.,
               Lehoczyk,J., Macdonald,P., Marquis,N., McEwan,P., McGurk,A.,
               McKernan,K., Meldrim,J., Mollia,M., Morris,W., Morrow,J.,
               Mychaleckyj,J., Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P.,
               Pavlin,B., Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A.,
               Severy,P., Stange-Thomann,N., Stojanovic,N., Stone,C.,
               Subramanian,A., Tesfaye,S., Torruella-Miller,I., Vassiliev,H.,
               Vo,A., Wagner,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
               Direct Submission
               Submitted (23-DEC-1998) Whitehead Institute/MIT Center for Genome
               Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT      On Dec 23, 1998 this sequence version replaced gi:4049327.
              All repeats were identified using RepeatMasker: Smit, A.F.A. &
              Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html.

FEATURES             source
                    1..191111
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="17"
                        /map="17"
                        /clone="hRPK.394_K10"
                        /clone_lib="RPCI-11 human BAC library"
                        457..574
                        /rpt_family="MIR"
                        complement(1149..1510)
                        /rpt_family="THE1B"
                        complement(1627..1996)
                        /rpt_family="THE1B"
                        2428..3054
                        /rpt_family="MER63"
                        3060..3127
                        /rpt_family="L2"
                        3147..3272
                        /rpt_family="MIR"
                        complement(4287..4365)
                        /note="Single-stranded terminator coverage."
                        4329..4398
                        /rpt_family="(TA)n"
                        complement(4541..4590)
                        /rpt_family="MLT1B"
                        complement(4591..4887)
                        /rpt_family="AluSg"
                        complement(4888..5231)
                        /rpt_family="MLT1B"
                        5232..5513
                        /rpt_family="AluYa5"
                        complement(5514..5553)
                        /rpt_family="MLT1B"
                        5570..5909
                        /rpt_family="L1ME1"
                        complement(6069..6299)
                        /rpt_family="MIR"
                        6370..6435
                        /rpt_family="MADE1"
                        complement(7081..7370)
                        /rpt_family="AluSg"
                        7646..7692
                        /rpt_family="AT-rich"
                        8643..8726
                        /rpt_family="AluSp/q"
                        complement(9022..9239)
                        /rpt_family="MER58A"
                        9320..9413
                        /rpt_family="MIR"
                        10146..10442
                        /rpt_family="LTR16C"
                        10847..11158
                        /rpt_family="AluX"
                        11161..11198
                        /rpt_family="(GAAA)n"
                        11308..11591
                        /rpt_family="AluYb8"
                        11592..11623
                        /rpt_family="(TAAAA)n"
                        complement(12075..12169)
                        /rpt_family="MER69"
                        12172..12253
                        /rpt_family="L1PA4"
                        12254..12273
                        /rpt_family="AT-rich"
                        complement(12396..12440)
                        /rpt_family="MER69"
                        12892..12925

```

```

repeat_region      /rpt_family="L2"
complement(12926..13207)
/rpt_family="AluY"
repeat_region      13208..13277
/rpt_family="L2"
complement(13279..13479)
/rpt_family="MIR"
repeat_region      13485..13580
/rpt_family="Tigter1"
complement(13581..13665)
/rpt_family="purine-rich"
repeat_region      complement(13666..13967)
/rpt_family="AluSx"
repeat_region      13969..15624
/rpt_family="Tigter1"
repeat_region      complement(15626..15792)
/rpt_family="L1MA5"
repeat_region      15793..16105
/rpt_family="AluSx"
repeat_region      complement(16106..16174)
/rpt_family="L1MA5"
repeat_region      16175..16514
/rpt_family="MER1B"
repeat_region      complement(16519..17773)
/rpt_family="L1MA5"
repeat_region      17776..18329
/rpt_family="Tigter1"
repeat_region      18719..18822
/rpt_family="AT rich"
repeat_region      complement(19223..20472)
/rpt_family="MER58B"
repeat_region      complement(20473..20762)
/rpt_family="AluUb"
repeat_region      complement(20763..20796)
/rpt_family="MER58B"
repeat_region      20947..21070
/rpt_family="L1ME3A"
repeat_region      21161..21179
/rpt_family="AT rich"
repeat_region      complement(21180..21478)
/rpt_family="AluSx"
repeat_region      21479..21505
/rpt_family="AT rich"
repeat_region      21556..21579
/rpt_family="TA)n"
repeat_region      complement(21580..21887)
/rpt_family="AluSx"
repeat_region      22010..22216
/rpt_family="MIR"
repeat_region      complement(22209..22252)
/rpt_family="MIR"
repeat_region      22502..22793
/rpt_family="AluSg"
repeat_region      24855..24911
/rpt_family="TA)n"
repeat_region      25813..26017
/rpt_family="L1M4c"
repeat_region      26018..26098
/rpt_family="purine-rich"
repeat_region      26127..26351
/rpt_family="purine-rich"
repeat_region      complement(26843..26866)
/rpt_family="POLY A"
repeat_region      27022..27325
/rpt_family="AluSx"
repeat_region      27960..28248
/rpt_family="AluSx"
repeat_region      30101..30179
/rpt_family="MIR"
repeat_region      complement(30335..30803)
/rpt_family="L1PA9"
repeat_region      complement(31111..31308)
/rpt_family="MTIA1"

```

```

repeat_region      complement(31358..31521)
/rpt_family="MLT1J"
repeat_region      32402..32706
/rpt_family="AluSx"

Query Match      55.2%; Score 23.2; DB 9; Length 191111;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY      3 ACCTACCCCGTGGTCACAAATGACCATGCTTGAC 38
Db 111225 ACATTCCCTGGAGGCAAAATGACCTGCTTGAC 111190

RESULT 14
AC114359/c
LOCUS      AC114359
DEFINITION Rattus norvegicus clone CH230-72G23, *** SEQUENCING IN PROGRESS
***, 74 unordered pieces.
ACCESSION  AC114359
VERSION     AC114359.3 GI:21738666
KEYWORDS    HTG: HTGS PHASE1.
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 194914)
AUTHORS    Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
            Alsbrooks, S.L., Amaratunga, H.C., Are, J.R., Ayale, M., Banks, I.,
            Barbara, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,
            Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
            Buhay, C., Burck, P., Burkett, C., Surrall, K.L., Byrd, N.C.,
            Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
            Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
            Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
            Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
            Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
            Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
            Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
            Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Prantz, P.,
            Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
            Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
            Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
            Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
            Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
            Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
            Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
            Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
            Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,
            Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
            Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
            Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
            Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
            Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
            Nguyen, N., Nickerson, E., Nwokkwo, S., Ogih, M., Okwundu, G.,
            Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
            Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
            Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
            Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
            Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,
            Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
            Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
            Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
            Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
            Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
            Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
            Weinstein, G., and Gibbs, R.
            Direct Submission
            Unpublished
            2 (bases 1 to 194914)
            Worley, K.C.
            Direct Submission
            TITLE
            JOURNAL
            REFERENCE
            AUTHORS
            TITLE

```

JOURNAL

Submitted (08-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 194914)

REFERENCE

Worley, K.C.

AUTHORS

Direct Submission

TITLE

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

JOURNAL

COMMENT

On Jul 12, 2002 this sequence version replaced gi:20303185.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GRPE

Center clone name: CH230-72G23

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990129

Consensus quality: 114397 bases at least Q40

Consensus quality: 119665 bases at least Q30

Consensus quality: 124866 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 74 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 1063: contig of 1063 bp in length

* 1064 1163: gap of unknown length

* 1164 2236: contig of 1073 bp in length

* 2237 2336: gap of unknown length

* 2337 3754: contig of 1418 bp in length

* 3755 3854: gap of unknown length

* 3855 5033: contig of 1179 bp in length

* 5034 5133: gap of unknown length

* 5134 6709: contig of 1576 bp in length

* 6710 6809: gap of unknown length

* 6810 8223: contig of 1414 bp in length

* 8224 8323: gap of unknown length

* 8324 9381: contig of 1038 bp in length

* 9382 9481: gap of unknown length

* 9482 11070: contig of 1589 bp in length

* 11071 11170: gap of unknown length

* 11171 12286: contig of 1116 bp in length

* 12287 12387: gap of unknown length

* 12387 13471: contig of 1085 bp in length

* 13472 13571: gap of unknown length

* 13572 15030: contig of 1459 bp in length

* 15031 15130: gap of unknown length

* 15131 16253: contig of 1123 bp in length

* 16254 16354: gap of unknown length

* 16354 17431: contig of 1078 bp in length

* 17432 17531: gap of unknown length

* 17532 18871: contig of 1340 bp in length

* 18872 18971: gap of unknown length

* 18972 20729: contig of 1758 bp in length

* 20730 20829: gap of unknown length

* 20830 22175: contig of 1346 bp in length

* 22176 22275: gap of unknown length

* 22276 23419: contig of 1144 bp in length

* 23420 23519: gap of unknown length

* 23520 24837: contig of 1318 bp in length

* 24838 24937: gap of unknown length

* 24938 26927: contig of 1990 bp in length

26928 27027: gap of unknown length

27028 28218: contig of 1191 bp in length

28219 28318: gap of unknown length

28319 29840: contig of 1522 bp in length

29841 29940: gap of unknown length

29941 31487: contig of 1547 bp in length

31488 31587: gap of unknown length

31588 32721: contig of 1134 bp in length

32722 32821: gap of unknown length

32822 34362: contig of 1541 bp in length

34363 34462: gap of unknown length

34463 36010: contig of 1548 bp in length

36011 36110: gap of unknown length

36111 37551: contig of 1441 bp in length

37552 37651: gap of unknown length

37652 39337: contig of 1686 bp in length

39338 39437: gap of unknown length

39438 41382: contig of 1945 bp in length

41383 41482: gap of unknown length

41483 43353: contig of 1871 bp in length

43354 43453: gap of unknown length

43454 45095: contig of 1642 bp in length

45096 45195: gap of unknown length

45196 46932: contig of 1737 bp in length

46933 47032: gap of unknown length

47033 48655: contig of 1623 bp in length

48656 48755: gap of unknown length

48756 50616: contig of 1861 bp in length

50617 52375: contig of 1659 bp in length

52376 52475: gap of unknown length

52476 54880: contig of 2405 bp in length

54881 54980: gap of unknown length

54981 56347: contig of 1367 bp in length

56348 56447: gap of unknown length

56448 58117: contig of 1670 bp in length

58118 58217: gap of unknown length

58218 60085: contig of 1868 bp in length

60086 60185: gap of unknown length

60186 62429: contig of 2244 bp in length

62430 62529: gap of unknown length

62530 64728: contig of 2199 bp in length

64729 64828: gap of unknown length

64829 67704: contig of 2876 bp in length

67705 67804: gap of unknown length

67805 69287: contig of 1483 bp in length

69288 69387: gap of unknown length

69388 71869: contig of 2482 bp in length

71870 71969: gap of unknown length

71970 75039: contig of 3070 bp in length

75040 75139: gap of unknown length

75140 77288: contig of 2149 bp in length

77289 77388: gap of unknown length

77389 81058: contig of 3670 bp in length

81059 81158: gap of unknown length

81159 82812: contig of 1654 bp in length

82813 82912: gap of unknown length

82913 84946: contig of 2034 bp in length

84947 85046: gap of unknown length

85047 88186: contig of 3140 bp in length

88187 88286: gap of unknown length

88287 90955: contig of 2669 bp in length

90956 91055: gap of unknown length

91056 93257: contig of 2202 bp in length

93258 93357: gap of unknown length

93358 96344: contig of 2987 bp in length

96345 96444: gap of unknown length

96445 98430: contig of 1986 bp in length

98431 98530: gap of unknown length

98531 101590: contig of 3060 bp in length

Query Match Score 23.2; DB 2; Length 194914;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;

Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 6 TACCCCGTGGGACAAATGACCATGGTTGACACA 41
 |||||
 Db 108304 TACCCCGTGGGACCAAGATGACCATGGAAACACA 108269
 |||||

RESULT 15

HSU63415

LOCUS

DEFINITION

Human peroxisome proliferator activated receptor gamma 2 mRNA, linear PRI 18-JUL-1996

ACCESSION U63415

VERSION U63415.1

KEYWORDS GI:1432176

SOURCE

ORGANISM

Homo sapiens.

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1518)

Moller, D.E. and Berger, J.

Molecular cloning, expression and characterization of human

peroxisome proliferator activated receptors gamma 1 and gamma 2

Biochem Biophys Res Commun. (1996) In press

2 (bases 1 to 1518)

Eibrecht, A., Chen, Y., Cullinan, C.A., Hayes, N., Leibowitz, M.D.,

Moller, D.E. and Berger, J.

Direct Submission

Submitted (10-JUL-1996) Molecular Endocrinology, merck Research

Labs., P.O. Box 2000, Mail Drop R80Y-265, Rahway, NJ 07060, USA

Location/Qualifiers

1. .1518

/organism="Homo sapiens"

/db_xref="taxon:9606"

/cell_type="fat cell"

1. .1518

/codon_start=1

/product="peroxisome proliferator activated receptor gamma

2"

/protein_id="AA04138.1"

/db_xref="GI:1432177"

/translation="MGETLG PDPDSSTDTLSANISOEMTMDTMDPMPNPG

ISSVLSWEDHSHSPDIK FTVDFSSISPHYEDIPFTRDPVVDYKYLKQEY

QSAIKVEPSPYPTQLYNKPHEEPSNLSMAIECRVCGDKASGFHYGACGCK

GFRTRILKLIYDRCLNCRIRHKRNKQYCFQKCLAVGMSHNAIRFGMPQAEK

EKLAEISSDIDQLNPESADLRALAKHLYDSYIKSFPLTKAKARAILTKGTTDKSPV

IYDMSLMMGEDKI KFKHITPQEQSKVAIRI FQCGFRSVEAOEITEYAKSIPGF

VNLNDQVTLKYGVHEIITVTLASLMNKOGVLISSGQGMTEBELKSLRKPQDEM

BKPFEFAKFNALFDDSLAIFIAVILSGDRFGLLNKVFIEDIODNLLQALEQLK

LNHPESOLFALQKMDLRQIVTEHVOLLQVVKTKTETDMSLHPLQEIYKDY"

BASE COUNT 438 a 363 c 346 g 371 t

ORIGIN

Query Match 54.8%; Score 23; DB 9; Length 1518;

Best Local Similarity 100.0%; Pred No. 2e+02;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AGAAATGACCATGGTTGACACAG 42

|||||

Db 81 AGAAATGACCATGGTTGACACAG 103

|||||

Search completed: January 21, 2003, 23:53:11

Job time : 1029.33 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 23:00:10 ; Search time 127.667 Seconds
(without alignments)
740.866 Million cell updates/sec

Title: US-09-765-111A-7

Perfect score: 42

Sequence: 1 agacctaccgcggtgtgca.....aatgacatggtgacacag 42

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002.*

1: /SID82/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID82/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID82/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID82/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID82/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID82/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID82/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID82/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID82/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID82/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID82/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID82/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID82/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID82/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID82/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID82/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID82/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID82/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID82/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID82/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID82/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID82/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID82/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID82/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	42	100.0	42	22	AAH76284
2	42	100.0	2334	22	AAH76281
3	30.4	72.4	42	22	AAH76286
4	30.4	72.4	2596	22	AAH76294
5	30.4	72.4	2625	22	AAH76283
6	30.2	71.9	1811	20	AAH76283
7	30.2	71.9	1811	22	AAH76296
8	30.2	71.9	1811	22	AAH76296
9	30.2	71.9	1844	22	AAH76296

c	10	30.2	71.9	2295	22	AAS44653	Human full-length
	11	25.2	60.0	42	22	AAH76285	PAX8e8-PPARGammael
	12	25.2	60.0	2523	22	AAH76282	Human PAX8e8-PPARG
	13	24.6	58.6	33	21	AA257563	Human PPAR gamma p
	14	24.6	58.6	33	22	AAH26111	Human PPAR gamma g
	15	24.6	58.6	33	22	AAH41855	Human PPAR gamma p
	16	24.6	58.6	33	22	AAH42011	Body weight gain i
	17	24	57.1	695	20	AAH19066	Human PPAR-gamma i
	18	23	54.8	26	20	AAH00271	Human PPAR-gamma p
	19	23	54.8	1518	22	AAH21022	Human peroxisome p
	20	23	54.8	1608	22	AAH76288	Human PPARgamma cd
	21	23	54.8	1608	22	AAH76288	Human PPARgamma co
	22	23	54.8	1647	17	AAT35334	Peroxisome prolif
	23	22	52.4	346	24	ABL37401	Human colon tumor
	24	22	52.4	611	21	AAC98099	Human colon cancer
	25	22	52.4	683	24	ABL83283	Human ovarian can
	26	22	52.4	1372	22	AAH76287	Human PAX8 cDNA se
	27	22	52.4	1936	17	AAT35333	Peroxisome prolif
	28	22	52.4	2757	22	AAH16672	Human cDNA sequen
	29	22	52.4	2757	22	AAH16672	Human cDNA sequen
	30	21.6	51.4	167	22	ABA70657	Human foetal liver
	31	21.6	51.4	167	22	AAK18903	Human brain expres
	32	21.6	51.4	167	22	AAK44847	Human bone marrow
	33	21.6	51.4	167	22	AAI50823	Probe #19509 used
	34	21.6	51.4	167	24	ABS19088	Human genome-deriv
	35	21.6	51.4	455	22	ABA58053	Human foetal liver
	36	21.6	51.4	455	22	AAK06125	Human brain expres
	37	21.6	51.4	455	22	AAK31778	Human bone marrow
	38	21.6	51.4	455	22	AAI37653	Probe #6339 used t
	39	21.6	51.4	455	24	ABS06536	Human genome-deriv
	40	21.6	51.4	2961	24	ABK83987	Human cDNA differe
	41	21.4	51.0	239	22	AAF81676	Bovine PPAR coding
	42	21.4	51.0	277	19	AAV21208	Rat PPAR-gamma A/B
	43	21.4	51.0	373	19	AAV21207	Mouse PPAR-gamma A
	44	21.2	50.5	277	15	AAQ76830	Human genome fragm
	45	21.2	50.5	569	23	AAH75424	DNA encoding novel

ALIGNMENTS

RESULT 1	AAH76284	ID	AAH76284	standard; DNA; 42 BP.
AC	AAH76284;	AC	AAH76284;	
XX		XX		
DT	29-OCT-2001	(first entry)		
XX				
DE	PAX8e7-PPARGammael translocation fusion juncture 1 DNA sequence.			
XX				
KW	PAX8-PPARGammael; oncogene; cytostatic; PAX8; PPARGammael; cancer;			
KW	follicular carcinoma; PAX8e7-PPARGammael; human; ss.			
XX				
OS	Homo sapiens.			
XX				
Key	Location/Qualifiers			
FT	3..41			
FT	/*tag= a			
XX				
PN	WO200152789-A2.			
XX				
PD	26-JUL-2001.			
XX				
PF	18-JAN-2001; 2001WO-US01664.			
XX				
PR	20-JAN-2000; 2000US-0177109.			
XX	14-AUG-2000; 2000US-0225079.			
XX	(BGHM) BRIGHAM & WOMENS HOSPITAL INC.			
PA				
XX				
PI	Kroll TG, Fletcher JA;			
XX				

DR WPI: 2001-514487/56.
 DR P-PSDB; AAB85796.
 PT New PAX8-PPARGmal oncogene and oncoprotein, useful for detecting and
 PT treating certain tumors or cancers, e.g. follicular carcinoma -
 XX
 PS Claim 21; Page 118; 145pp; English.
 XX
 CC The invention relates to an oncogene designated PAX8-PPARGmal that
 CC contains a PAX8 coding region fused to PPARGmal coding region. The
 CC PAX8-PPARGmal polypeptides can be expressed by standard recombinant
 CC methodology. A PPARGmal ligand or agent is useful for treating a
 CC subject having a disorder characterized by the presence of a PAX8-
 CC PPARGmal, where the disorder is cancer, e.g. follicular carcinoma.
 CC The PAX8-PPARGmal molecules are also useful for providing nucleotide
 CC and amino acid sequences useful for detecting the above disease. The
 CC present sequence represents the nucleotide sequence including and
 CC surrounding the translocation fusion junction in the PAX8e7-PPARGmal
 CC DNA.
 XX
 SQ Sequence 42 BP; 12 A; 11 C; 12 G; 7 T; 0 other;
 Query Match 100.0%; Score 42; DB 22; Length 42;
 Best Local Similarity 100.0%; Pred. No. 9e-08;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 AC AAH76281;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Human PAX8e7-PPARGmal cDNA sequence.
 XX
 KW PAX8-PPARGmal; oncogene; cytostatic; PAX8; PPARGmal; cancer;
 KW follicular carcinoma; PAX8e7-PPARGmal; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2334
 FT /*tag= a
 XX
 PN WO200152789-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 18-JAN-2001; 2001WO-US01664.
 XX
 PR 20-JAN-2000; 2000US-0177109.
 PR 14-AUG-2000; 2000US-0225079.
 XX
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
 XX
 PI Kroll TG, Fletcher JA;
 XX
 DR WPI: 2001-514487/56.
 DR P-PSDB; AAB85796.
 XX
 PT New PAX8-PPARGmal oncogene and oncoprotein, useful for detecting and
 PT treating certain tumors or cancers, e.g. follicular carcinoma -
 XX
 PS Claim 1; Page 100-104; 145pp; English.
 XX
 CC The invention relates to an oncogene designated PAX8-PPARGmal that
 CC contains a PAX8 coding region fused to PPARGmal coding region. The

CC PAX8-PPARGmal polypeptides can be expressed by standard recombinant
 CC methodology. A PPARGmal ligand or agent is useful for treating a
 CC subject having a disorder characterized by the presence of a PAX8-
 CC PPARGmal, where the disorder is cancer, e.g. follicular carcinoma.
 CC The PAX8-PPARGmal molecules are also useful for providing nucleotide
 CC and amino acid sequences useful for detecting the above disease. The
 CC present sequence represents a human PAX8e7-PPARGmal polypeptide
 CC encoding cDNA.
 XX
 SQ Sequence 2334 BP; 614 A; 650 C; 573 G; 497 T; 0 other;
 Query Match 100.0%; Score 42; DB 22; Length 2334;
 Best Local Similarity 100.0%; Pred. No. 2.1e-07;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGACCTACCCCGTGGCGAGAAATGACCATGTTGACACAG 42
 DB 878 AGACCTACCCCGTGGCGAGAAATGACCATGTTGACACAG 919
 RESULT 3
 AAH76286
 ID AAH76286 standard; DNA; 42 BP.
 XX
 AC AAH76286;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE PAX8e9-PPARGmal fusion junction DNA sequence.
 XX
 KW PAX8-PPARGmal; oncogene; cytostatic; PAX8; PPARGmal; cancer;
 KW follicular carcinoma; PAX8e9-PPARGmal; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 3..41
 FT /*tag= a
 XX
 PN WO200152789-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 18-JAN-2001; 2001WO-US01664.
 XX
 PR 20-JAN-2000; 2000US-0177109.
 PR 14-AUG-2000; 2000US-0225079.
 XX
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
 XX
 PI Kroll TG, Fletcher JA;
 XX
 DR WPI: 2001-514487/56.
 DR P-PSDB; AAB85796.
 XX
 PT New PAX8-PPARGmal oncogene and oncoprotein, useful for detecting and
 PT treating certain tumors or cancers, e.g. follicular carcinoma -
 XX
 PS Claim 21; Page 119; 145pp; English.
 XX
 CC The invention relates to an oncogene designated PAX8-PPARGmal that
 CC contains a PAX8 coding region fused to PPARGmal coding region. The
 CC PAX8-PPARGmal polypeptides can be expressed by standard recombinant
 CC methodology. A PPARGmal ligand or agent is useful for treating a
 CC subject having a disorder characterized by the presence of a PAX8-
 CC PPARGmal, where the disorder is cancer, e.g. follicular carcinoma.
 CC The PAX8-PPARGmal molecules are also useful for providing nucleotide
 CC and amino acid sequences useful for detecting the above disease. The
 CC present sequence represents the nucleotide sequence including and
 CC surrounding the fusion junction in the PAX8e9-PPARGmal DNA.
 XX
 SQ Sequence 42 BP; 12 A; 10 C; 13 G; 7 T; 0 other;

Query Match 72.4%; Score 30.4; DB 22; Length 42;
 Best Local Similarity 96.9%; Pred. No. 0.0052;
 Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 11 CQTGTCGCAGAAATGACCATGGTTGACACAG 42
 DB 11 CATGTCGCAGAAATGACCATGGTTGACACAG 42

RESULT 4
 AAH76294
 ID AAH76294 standard; cDNA; 2596 BP.
 AC AAH76294;
 XX
 DT 29-OCT-2001 (first entry)
 DE Human PAX8e9(-exon 8)-PPARGammael cDNA sequence.
 XX
 KW PAX8-PPARGammael; oncogene; cytostatic; PAX8; PPARGammael; cancer;
 KW follicular carcinoma; PAX8e9(-exon 8)-PPARGammael; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 161..2596
 FT /*tag= a
 XX
 PN WO200152789-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 18-JAN-2001; 2001WO-US01664.
 XX
 PR 20-JAN-2000; 2000US-0177109.
 PR 14-AUG-2000; 2000US-0225079.
 XX
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
 XX
 PI Kroll TG, Fletcher JA;
 XX
 DR WPI; 2001-514487/56.
 DR P-PSDB; AAB85801.
 XX
 XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
 P. creating certain tumors or cancers, e.g. follicular carcinoma
 PS Claim 1; Page 127-131; 145pp; English.
 XX
 CC The invention relates to an oncogene designated PAX8-PPARGammael that
 CC contains a PAX8 coding region fused to PPARGammael coding region. The
 CC PAX8-PPARGammael polypeptides can be expressed by standard recombinant
 CC methodology. A PPARGammael ligand or agent is useful for treating a
 CC subject having a disorder characterized by the presence of a PAX8-
 CC PPARGammael, where the disorder is cancer, e.g. follicular carcinoma.
 CC The PAX8-PPARGammael molecules are also useful for providing nucleotide
 CC and amino acid sequences useful for detecting the above disease. The
 CC present sequence represents a human PAX8e9(-exon 8)-PPARGammael
 XX polypeptide encoding cDNA.
 SQ Sequence 2596 BP; 660 A; 745 C; 672 G; 519 T; 0 other;

Query Match 72.4%; Score 30.4; DB 22; Length 2596;
 Best Local Similarity 96.9%; Pred. No. 0.013;
 Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 11 CQTGTCGCAGAAATGACCATGGTTGACACAG 42
 DB 1150 CATGTCGCAGAAATGACCATGGTTGACACAG 1181

RESULT 5
 AAH76283
 ID AAH76283 standard; cDNA; 2625 BP.
 AC AAH76283;
 XX
 DT 29-OCT-2001 (first entry)
 DE Human PAX8e9-PPARGammael cDNA sequence.
 XX
 KW PAX8-PPARGammael; oncogene; cytostatic; PAX8; PPARGammael; cancer;
 KW follicular carcinoma; PAX8e9-PPARGammael; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2625
 FT /*tag= a
 XX
 PN WO200152789-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 18-JAN-2001; 2001WO-US01664.
 XX
 PR 20-JAN-2000; 2000US-0177109.
 PR 14-AUG-2000; 2000US-0225079.
 XX
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
 XX
 PI Kroll TG, Fletcher JA;
 XX
 DR WPI; 2001-514487/56.
 DR P-PSDB; AAB85795.
 XX
 XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
 PT treating certain tumors or cancers, e.g. follicular carcinoma
 PS Claim 1; Page 112-116; 145pp; English.
 XX
 CC The invention relates to an oncogene designated PAX8-PPARGammael that
 CC contains a PAX8 coding region fused to PPARGammael coding region. The
 CC PAX8-PPARGammael polypeptides can be expressed by standard recombinant
 CC methodology. A PPARGammael ligand or agent is useful for treating a
 CC subject having a disorder characterized by the presence of a PAX8-
 CC PPARGammael, where the disorder is cancer, e.g. follicular carcinoma.
 CC The PAX8-PPARGammael molecules are also useful for providing nucleotide
 CC and amino acid sequences useful for detecting the above disease. The
 CC present sequence represents a human PAX8e9-PPARGammael polypeptide
 XX encoding cDNA.
 SQ Sequence 2625 BP; 662 A; 761 C; 646 G; 556 T; 0 other;

Query Match 72.4%; Score 30.4; DB 22; Length 2625;
 Best Local Similarity 96.9%; Pred. No. 0.013;
 Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 11 CQTGTCGCAGAAATGACCATGGTTGACACAG 42
 DB 1179 CATGTCGCAGAAATGACCATGGTTGACACAG 1210

RESULT 6
 AAH76282
 ID AAH76282 standard; cDNA; 1811 BP.
 AC AAH76282;
 XX
 DT 07-JUL-1999 (first entry)
 DE Human PPARG-gammael coding sequence.
 XX
 KW Nuclear receptor agonist; antagonist; identification; PPARG;
 KW peroxisome proliferator activated receptor; ss.
 XX

OS Homo sapiens.
 PN WO9918124-A1.
 PD 15-APR-1999.
 PF 06-OCT-1998; 98WO-US21049.
 PR 07-OCT-1997; 97US-0061385.
 PA (MERI) MERCK & CO INC.
 PI Cummings RT, Hermes JD, Moller DE, Zhou G;
 DR WPI; 1999-263998/22.
 DR P-PSDB; AAY05471.
 XX
 PT Identifying nuclear receptor agonists and antagonists
 PS Disclosure; Fig 9b; 60pp; English.
 CC This sequence encodes the human peroxisome proliferator activated
 CC receptor-gamma1 (PPAR-gamma1).
 CC The invention relates to a method for identifying nuclear receptor
 CC agonists and antagonists comprises measuring fluorescent resonance energy
 CC transfer between fluorescent-labelled nuclear receptors and
 CC co-activators. The method can be used for identifying agonists and
 CC antagonist of nuclear receptors.
 XX
 SQ Sequence 1811 BP; 510 A; 433 C; 422 G; 446 T; 0 other;
 Query Match 71.9%; Score 30.2; DB 20; Length 1811;
 Best Local Similarity 91.4%; Pred. No. 0.014;
 Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 8 CCCGCTGGTGGCAGAAATGACCATGGTTGACACAG 42
 DB 157 CCCGCTGGCAGAAATGACCATGGTTGACACAG 191
 RESULT 7
 AAH76296
 ID AAH76296 standard; cDNA; 1811 BP.
 XX
 AC AAH76296;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Human PPARgamma cDNA sequence.
 XX
 KW PAX8-PPARGamma1; oncogene; cytostatic; PAX8; PPARGamma1; cancer;
 KW follicular carcinoma; PPARGamma; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 173..1609
 FT /*tag= a
 XX
 XX WO200152789-A2.
 PN
 PD 26-JUL-2001.
 XX
 PF 18-JAN-2001; 2001WO-US01664.
 XX
 PR 20-JAN-2000; 2000US-0177109.
 PR 14-AUG-2000; 2000US-0225079.
 XX
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
 XX
 PI Kroll TG, Fletcher JA;
 XX
 WPI; 2001-514487/56.

DR P-PSDB; AAB85802.
 XX
 PT New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
 PT treating certain tumors or cancers, e.g. follicular carcinoma
 XX
 PS Disclosure; Page 137-139; 145pp; English.
 XX
 CC The invention relates to an oncogene designated PAX8-PPARGamma1 that
 CC contains a PAX8 coding region fused to PPARGamma1 coding region. The
 CC PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant
 CC methodology. A PPARGamma ligand or agent is useful for treating a
 CC subject having a disorder characterized by the presence of a PAX8-
 CC PPARGamma1, where the disorder is cancer, e.g. follicular carcinoma.
 CC The PAX8-PPARGamma1 molecules are also useful for providing nucleotide
 CC and amino acid sequences useful for detecting the above disease. The
 CC present sequence represents a human PPARGamma polypeptide encoding cDNA.
 XX
 SQ Sequence 1811 BP; 510 A; 433 C; 422 G; 446 T; 0 other;
 Query Match 71.9%; Score 30.2; DB 22; Length 1811;
 Best Local Similarity 91.4%; Pred. No. 0.014;
 Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 8 CCCGCTGGTGGCAGAAATGACCATGGTTGACACAG 42
 DB 157 CCCGCTGGCAGAAATGACCATGGTTGACACAG 191
 RESULT 8
 ABN95714
 ID ABN95714 standard; DNA; 1811 BP.
 XX
 AC ABN95714;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Gene #2212 used to diagnose liver cancer.
 XX
 KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 KW metastatic liver tumour; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX
 OS Homo sapiens.
 XX
 PN WO200229103-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 02-OCT-2001; 2001WO-US30589.
 XX
 PR 02-OCT-2000; 2000US-237054P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 XX
 WPI; 2002-426119/45.
 DR
 XX
 PT Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample
 XX
 PS Claim 1; SEQ ID NO 2212; 298pp; English.
 XX
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic

CC in culture to give rise to neuroepithelial cells that can be used to
 CC augment or replace cells damaged by illness, accidental damage or genetic
 CC disorders. The sequences may also be used for regeneration of bone,
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.
 CC Note: Some sequences for this patent did not form part of the printed
 CC specification but were obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.

XX SQ Sequence 2295 BP; 525 A; 606 C; 565 G; 599 T; 0 other;

Query Match 71.9%; Score 30.2; DB 22; Length 2295;
 Best Local Similarity 91.4%; Pred. No. 0.015;
 Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCCTGGTGGCAGAAATGACCATGGTTGACACAG 42
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1655 CGCGTGGCGCAGAAATGACCATGGTTGACACAG 1621

RESULT 11
 AAH76285
 ID AAH76285 standard; DNA; 42 BP.

AC AAH76285;
 XX 29-OCT-2001 (first entry)

DE PAX8e8-PPARGammael fusion junction DNA sequence.

XX PAX8-PPARGammael; oncogene; cytostatic; PAX8; PPARGammael; cancer;
 KW follicular carcinoma; PAX8e8-PPARGammael; human; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 3..41
 FT /*tag= a

PN WO200152789-A2.

PD 26-JUL-2001.

PF 18-JAN-2001; 2001WO-US01664.

XX 20-JAN-2000; 2000US-0177109.

PR 14-AUG-2000; 2000US-0225079.

PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.

XX Kroll TG, Fletcher JA;

XX WPI; 2001-514487/56.

DR P-PSDB; AAB85797.

XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
 PT treating certain tumors of cancers, e.g. follicular carcinoma

XX Claim 21; Page 118; 145pp; English.

XX The invention relates to an oncogene designated PAX8-PPARGammael that
 CC contains a PAX8 coding region fused to PPARGammael coding region. The
 CC PAX8-PPARGammael polypeptides can be expressed by standard recombinant
 CC methodology. A PPARGammael ligand or agent is useful for treating a
 CC subject having a disorder characterized by the presence of a PAX8-
 CC PPARGammael, where the disorder is cancer, e.g. follicular carcinoma.
 CC The PAX8-PPARGammael molecules are also useful for providing nucleotide
 CC and amino acid sequences useful for detecting the above disease. The
 CC present sequence represents the nucleotide sequence including and
 CC surrounding the fusion junction in the PAX8e8-PPARGammael DNA.

XX SQ Sequence 42 BP; 10 A; 14 C; 11 G; 7 T; 0 other;
 Query Match 60.0%; Score 25.2; DB 22; Length 42;

Best Local Similarity 78.9%; Pred. No. 0.7;
 Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 CTACCCCGTGGTGGCAGAAATGACCATGGTTGACACAG 42
 | | | | | | | | | | | | | | | | | | | | | |
 Db 5 CCAGGCCCTCCTCTCAGAAATGACCATGGTTGACACAG 42

RESULT 12
 AAH76282
 ID AAH76282 standard; cDNA; 2523 BP.

XX AC AAH76282;

XX 29-OCT-2001 (first entry)

XX Human PAX8e8-PPARGammael cDNA sequence.

XX PAX8-PPARGammael; oncogene; cytostatic; PAX8; PPARGammael; cancer;
 KW follicular carcinoma; PAX8e8-PPARGammael; human; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 1..2523
 FT /*tag= a

PN WO200152789-A2.

PD 26-JUL-2001.

PF 18-JAN-2001; 2001WO-US01664.

XX 20-JAN-2000; 2000US-0177109.

PR 14-AUG-2000; 2000US-0225079.

PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.

XX Kroll TG, Fletcher JA;

XX WPI; 2001-514487/56.

DR P-PSDB; AAB85794.

XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
 PT treating certain tumors or cancers, e.g. follicular carcinoma

XX Claim 1; Page 106-109; 145pp; English.

XX The invention relates to an oncogene designated PAX8-PPARGammael that
 CC contains a PAX8 coding region fused to PPARGammael coding region. The
 CC PAX8-PPARGammael polypeptides can be expressed by standard recombinant
 CC methodology. A PPARGammael ligand or agent is useful for treating a
 CC subject having a disorder characterized by the presence of a PAX8-
 CC PPARGammael, where the disorder is cancer, e.g. follicular carcinoma.
 CC The PAX8-PPARGammael molecules are also useful for providing nucleotide
 CC and amino acid sequences useful for detecting the above disease. The
 CC present sequence represents a human PAX8e8-PPARGammael polypeptide
 CC encoding cDNA.

XX SQ Sequence 2523 BP; 643 A; 723 C; 614 G; 543 T; 0 other;

Query Match 60.0%; Score 25.2; DB 22; Length 2523;
 Best Local Similarity 78.9%; Pred. No. 1.7;
 Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 CTACCCCGTGGTGGCAGAAATGACCATGGTTGACACAG 42
 | | | | | | | | | | | | | | | | | | | | | |
 Db 1071 CCAGGCCCTCCTCTCAGAAATGACCATGGTTGACACAG 1108

RESULT 13
 AAZ57563
 ID AAZ57563 standard; DNA; 33 BP.

XX AC AA257563;
 XX DT 14-APR-2000 (first entry)
 XX DE Human PPAR gamma PCR primer PAG-U SEQ ID NO:1.
 XX KW Retinoid associated receptor regulator; 1,3-azole; diabetes;
 KW anti-diabetic; antilipemic; anorectic; nephrotic; osteopathic;
 KW anabolic; cytostatic; hypotensive; cardiac; anti-anginal; obesity;
 KW cerebroprotective; antiarteriosclerotic; hyperlipaemia; osteoporosis;
 KW cachexia; cancer; vascular disorder; hypertension; kidney disorder;
 KW muscular dystrophy; cardiac infarction; angina pectoris; syndrome X;
 KW cerebral infarction; hyperinsulinaemia; leukaemia; arteriosclerosis;
 KW PPAR gamma; PCR primer; ss.
 XX OS lomo sapiens.
 XX PN WO200001679-A1.
 XX PD 13-JAN-2000.
 XX PF 30-JUN-1999; 99WO-JP03520.
 XX PR 01-JUL-1998; 98JP-0186698.
 XX PA (TAKE) TAKEDA CHEM IND LTD.
 XX PI Sugiyama Y, Momose Y, Kimura H, Sakamoto J, Odaka H;
 XX WPI; 2000-137191/12.
 XX DR Retinoid associated receptor regulators, useful for treating diabetes,
 XX PT hyperlipemia and obesity
 XX PS Example; Page 58; 122pp; Japanese.
 XX CC The present invention describes retinoid associated receptor regulators
 CC comprising a 1,3-azole derivative (I). The retinoid associated receptor
 CC regulators are useful as insulin receptor binding agents for treating
 CC and preventing diabetes, hyperlipaemia and obesity. (I) may also be
 CC useful for the treatment and prevention of diabetic complications
 CC (e.g. nerve, kidney, retina, blood vessel and bone narrowing disorders
 CC and diseases), osteoporosis, cachexia (e.g. due to cancer, diabetes or
 CC vascular disorders), fatty liver, hypertension, polycystic ovary
 CC syndrome, kidney disorders (e.g. glomerulonephritis, diabetic
 CC nephropathy and glomerular sclerosis), muscular dystrophy, cardiac
 CC infarction, angina pectoris, cerebral infarction, disorders due to
 CC hyperinsulinaemia, syndrome X, cancer (e.g. leukaemia and breast cancer),
 CC and arteriosclerosis. The present sequence represents a PCR primer for
 CC human PPAR gamma, which is used in the exemplification of the present
 CC invention.
 XX SQ Sequence 33 BP; 10 A; 6 C; 11 G; 6 T; 0 other;
 Query Match 58.6%; Score 24.6; DB 21; Length 33;
 Best Local Similarity 87.1%; Pred. No. 1.2;
 Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 12 GTGGTGACGAAATGACCATGTTGACACAG 42
 DB 1 GTGGGTACCGAAATGACCATGTTGACACAG 31
 RESULT 14
 AAH26111
 ID AAH26111 standard; DNA; 33 BP.
 XX AC AAH26111;
 XX DT 17-SEP-2001 (first entry)
 XX DE Human PPAR-gamma gene PCR primer PAG-U.
 XX KW Human; PPAR gamma; RXR alpha; PPRE; herpes simplex virus; HSV;
 KW thymidine kinase minimum promoter; retinoid X receptor; anti-diabetic;
 KW peroxisome proliferation-activated receptor; neuroprotective; muscular;
 KW nephrotropic; ophthalmological; osteopathic; antilipemic; hypotensive;
 KW immunosuppressive; cytostatic; antibacterial; anti-HIV; endocrine;

XX KW PPAR-gamma; peroxisome proliferator-activated receptor gamma;
 KW human; hypoglycaemic; hypolipemic; anti-diabetic; anti-inflammatory;
 KW antiarteriosclerotic; diabetes; hyperlipidaemia;
 KW glucose intolerance; PCR primer; ss.
 XX OS Homo sapiens.
 XX PN WO200138325-A1.
 XX PD 31-MAY-2001.
 XX PF 09-NOV-2000; 2000MO-JP07877.
 XX PR 10-NOV-1999; 99JP-0320317.
 XX PR 10-DEC-1999; 99JP-0352237.
 XX PA (TAKE) TAKEDA CHEM IND LTD.
 XX PI Momose Y, Maekawa T, Odaka H, Kimura H;
 XX WPI; 2001-432602/46.
 XX DR 5-Membered N-heterocyclic derivatives useful for preventing or treating
 XX PT diabetes mellitus, hyperlipidaemia or impaired glucose tolerance
 XX PS Example 1; Page 85; 376pp; English.
 XX CC The present sequence is that of primer PAG-U, which was used with
 CC primer PAG-L (see AAH26112) in the PCR amplification of human
 CC peroxisome proliferator-activated receptor gamma (PPAR-gamma) cDNA,
 CC using heart cDNA as template. The primers are based on the
 CC reported base sequence of the PPAR-gamma gene. PPAR-gamma is a
 CC member of the intranuclear hormone receptor superfamily. Its
 CC expression is induced at a very early stage of adipose cell
 CC differentiation. PPAR-gamma forms a dimer with the retinoid X
 CC receptor (RXR) by binding to a ligand, and binds to a responsive
 CC site of a target gene in the nucleus to directly control (activate)
 CC transcription. PPAR-gamma ligands suppress the production of
 CC inflammatory cytokines by monocytes. The invention provides
 CC novel 5-membered N-heterocyclic compounds having hypoglycaemic and
 CC hypolipidaemic action, making them useful as agents for preventing
 CC or treating diabetes mellitus, hyperlipidaemia, impaired glucose
 CC tolerance, inflammatory disease and arteriosclerosis. The novel
 CC compounds have potent PPAR-gamma-RXR-alpha heterodimer ligand
 CC activity.
 XX SQ Sequence 33 BP; 10 A; 6 C; 11 G; 6 T; 0 other;
 Query Match 58.6%; Score 24.6; DB 22; Length 33;
 Best Local Similarity 87.1%; Pred. No. 1.2;
 Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 12 GTGGTGACGAAATGACCATGTTGACACAG 42
 DB 1 GTGGGTACCGAAATGACCATGTTGACACAG 31
 RESULT 15
 AAH41855
 ID AAH41855 standard; DNA; 33 BP.
 XX AC AAH41855;
 XX DT 30-AUG-2001 (first entry)
 XX DE Human PPAR gamma PCR primer PAG-U SEQ ID NO:1.
 XX KW Human; PPAR gamma; RXR alpha; PPRE; herpes simplex virus; HSV;
 KW thymidine kinase minimum promoter; retinoid X receptor; anti-diabetic;
 KW peroxisome proliferation-activated receptor; neuroprotective; muscular;
 KW nephrotropic; ophthalmological; osteopathic; antilipemic; hypotensive;
 KW immunosuppressive; cytostatic; antibacterial; anti-HIV; endocrine;

KW cardiant; cerebroprotective; antiinflammatory; antiarthritic; antiulcer;
 KW antineumatic; hepatotropic; respiratory; gastrointestinal; neuropathy;
 KW diabetic complication; neuropathy; retinopathy; osteopathy; obesity;
 KW hypertension; cachexia; cancer; infection; AIDS; renal disorder; tumour;
 KW polycystic ovary syndrome; muscular dystrophy; myocardial infarction;
 KW cerebral vascular disorder; syndrome X; inflammatory disease;
 KW arteriosclerosis; PCR primer; ss.

XX Homo sapiens.

XX WO200134579-A1.

XX 17-MAY-2001.

XX 09-NOV-2000; 2000WO-JP07878.

XX 10-NOV-1999; 99JP-0320318.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Momose Y, Imoto H, Odaka H, Kimura H;

XX WPI; 2001-355474/37.

XX Use of new and known alkoxyiminoalkanoic acid derivatives as
 PT peroxisome proliferation-activated receptor agonist like substances for
 PT treating e.g. diabetes -

XX Example; Page 66; 92pp; Japanese.

XX The present invention describes the use of new and known
 CC alkoxyiminoalkanoic acid derivatives (I). Alkoxyiminoalkanoic acid
 CC derivatives and their salts and prodrugs can be used for treating or
 CC preventing diabetes, hyperlipemia or impaired glucose tolerance or as a
 CC ligand for peroxisome proliferation-activated receptors (PPAR) or
 CC retinoid X receptors (RXR). (I) have antidiabetic, neuroprotective,
 CC nephrotropic, ophthalmological, osteopathic, antilipemic, hypotensive,
 CC muscular, immunosuppressive, cytostatic, antibacterial, anti-HIV,
 CC endocrine, cardiant, cerebroprotective, antiinflammatory, antiarthritic,
 CC antineumatic, hepatotropic, respiratory, antiulcer and gastrointestinal
 CC activities. (I) can be used for treating or preventing diabetic
 CC complications (such as neuropathy, nephropathy, retinopathy or
 CC osteopathy), obesity, hypertension, cachexia (e.g. due to cancer,
 CC infections or AIDS), polycystic ovary syndrome, renal disorders (e.g.
 CC glomerulonephritis), muscular dystrophy, myocardial infarction, cerebral
 CC vascular disorders (e.g. cerebral infarction or cerebral apoplexy),
 CC syndrome X, tumours, inflammatory diseases (e.g. chronic rheumatoid
 CC arthritis, hepatitis, pneumonia or ulcerative colitis) and
 CC arteriosclerosis. The present sequence represents a PCR primer for human
 CC PPAR gamma, which is used in an example from the present invention.

XX Sequence 33 BF; 10 A; 6 C; 11 G; 6 T; 0 other;

Query March 58.6%; Score 24.6; DB 22; Length 33;

Best Local Similarity 87.1%; Pred. No. 1.2;

Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 12 GTGGTGACAAATGACCATGTTGACACAG 42

Db GTGGTGACCAATGACCATGTTGACACAG 31

Search completed: January 21, 2003, 23:09:52
 Job time : 130.667 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 21, 2003, 23:03:25 ; Search time 26.3333 Seconds
(without alignments)
489.130 Million cell updates/sec

Title: US-09-765-111A-7
Perfect score: 42
Sequence: 1 agacctaccctgtgtgca.....aatgacctgtgtgacacag 42

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query Match	Length	ID	Description
1	30.2	71.9	1844	4	US-08-134-557D-1
2	23	54.8	1518	4	US-09-128-142-3
3	23	54.8	1608	3	US-09-484-345-3
4	23	54.8	1679	4	US-09-514-247A-5
5	21.4	51.0	277	3	US-08-917-653-4
6	21.4	51.0	373	3	US-08-917-653-3
7	21	50.0	10718	3	US-08-325-426B-1
8	20.8	49.5	1796	4	US-09-255-392-1
9	20.8	49.5	2005	2	US-08-484-200-1
10	20.8	49.5	2005	2	US-08-477-493-1
11	20.8	49.5	2005	3	US-08-465-375-1
12	20.8	49.5	2005	4	US-09-788-070-1
13	19.8	47.1	246240	2	US-08-724-394A-20
14	19.8	47.1	246240	2	US-08-724-394A-21
15	19.8	47.1	246240	2	US-08-724-394A-22
16	19.6	46.7	834	4	US-08-191-160-5
17	19.6	46.7	1107	4	US-09-000-094-19
18	19.6	46.7	1128	4	US-09-000-094-21
19	19.6	46.7	1398	4	US-09-000-094-23
20	19.6	46.7	2116	4	US-08-191-160-21
21	19.6	46.7	3621	4	US-09-220-081-1
22	19.6	46.7	3621	4	US-09-677-575-1
23	19.6	46.7	4770	4	US-09-000-094-45
24	19.4	46.2	738	4	US-09-182-145-38
25	19.4	46.2	841	4	US-09-182-145-39
26	19.4	46.2	1134	4	US-09-206-059-29
27	19.4	46.2	1293	4	US-09-182-145-13

28	19.4	46.2	1293	4	US-09-182-145-14
29	19.4	46.2	1784	3	US-09-226-741-2
30	19.4	46.2	1784	4	US-09-593-514-2
31	19.4	46.2	1982	3	US-09-226-741-4
32	19.4	46.2	1982	4	US-09-593-514-4
33	19.4	46.2	2296	1	US-07-750-080A-18
34	19.4	46.2	2296	3	US-08-651-472-18
35	19.4	46.2	2296	4	US-08-358-928-18
36	19.4	46.2	2497	1	US-08-643-219-12
37	19.4	46.2	2497	2	US-09-131-995-12
38	19.4	46.2	2497	2	US-08-832-087B-12
39	19.4	46.2	2497	3	US-08-851-350-12
40	19.4	46.2	2497	4	US-09-132-154-12
41	19.4	46.2	2679	6	5200340-7
42	19.4	46.2	2753	1	US-07-854-603-1
43	19.2	45.7	1964	2	US-08-841-349-8
44	19.2	45.7	1980	2	US-08-766-858A-4
45	19	45.2	1325	2	US-08-464-517-1

ALIGNMENTS

RESULT 1
US-08-134-557D-1
; Sequence 1, Application US/08134557D
; Patent No. 6200802
; GENERAL INFORMATION:
; APPLICANT: Greene, Marianne E.
; APPLICANT: Blumberg, Bruce
; TITLE OF INVENTION: Human Peroxisome Proliferator Activated
; RECEPTOR GAMMA: Compositions and Methods
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rockey, Milmanow & Katz, Ltd.
; STREET: 2 Prudential Plaza, Suite 4700 180 N. Stetson
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/134,557D
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Katz, Martin L.
; REGISTRATION NUMBER: 25,011
; REFERENCE/DOCKET NUMBER: ARCH:098
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5400
; TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1844 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 179..1606
; US-08-134-557D-1

Query Match 71.9%; Score 30.2; DB 4; Length 1844;
Best Local Similarity 91.4%; Pred. No. 0.0011;
Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 8 CCCCCTGTGTGCAGAAATGACCATGTTGACACAG 42

```
Db 157 CCGCGGCGCAGAAATGACCATGTTGACACAG 191
|||||
Query Match 54.8%; Score 23; DB 4; Length 1518;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 3
US-09-484-345-3
; Sequence 3, Application US/09128142
; Patent No. 6294559
; GENERAL INFORMATION:
; APPLICANT: Smith, Roy G.
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENTS ASSOCIATED WITH
; PEROXISOME PROLIFERATOR ACTIVATED RECEPTORS GAMMA1 AND GAMMA2
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jack L. Tribble
; STREET: 126 E. Lincoln Ave., P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128.142
; FILING DATE: 03-Aug-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,007
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Tribble, Jack L.
; REGISTRATION NUMBER: 32,633
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-5321
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1518 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (Genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-128-142-3

Query Match 54.8%; Score 23; DB 4; Length 1518;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AGAAATGACCATGTTGACACAG 42
|||||
Db 81 AGAAATGACCATGTTGACACAG 103
|||||

RESULT 3
US-09-484-345-3
; Sequence 3, Application US/09484345
; Patent No. 6159734
; GENERAL INFORMATION:
; APPLICANT: Robert McKay
; APPLICANT: Alexander H. Borchers
; APPLICANT: Brenda F. Baker
; TITLE OF INVENTION: ANTISENSE MODULATION OF PEROXISOME PROLIFERATOR-ACTIVATED RECEPTORS
; FILE REFERENCE: RTS-0104
; CURRENT APPLICATION NUMBER: US/09/484,345
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 3
; LENGTH: 1608

Db 157 CCGCGGCGCAGAAATGACCATGTTGACACAG 191
|||||
Query Match 54.8%; Score 23; DB 3; Length 1608;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2
US-09-128-142-3
; Sequence 3, Application US/09128142
; Patent No. 6294559
; GENERAL INFORMATION:
; APPLICANT: Smith, Roy G.
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENTS ASSOCIATED WITH
; PEROXISOME PROLIFERATOR ACTIVATED RECEPTORS GAMMA1 AND GAMMA2
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jack L. Tribble
; STREET: 126 E. Lincoln Ave., P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Power Macintosh 7500/100
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128.142
; FILING DATE: 03-Aug-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,007
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Tribble, Jack L.
; REGISTRATION NUMBER: 32,633
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-5321
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1518 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (Genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-128-142-3

Query Match 54.8%; Score 23; DB 4; Length 1608;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AGAAATGACCATGTTGACACAG 42
|||||
Db 239 AGAAATGACCATGTTGACACAG 261
|||||

RESULT 5
US-08-917-653-4
; Sequence 4, Application US/08917653
; Patent No. 6004751
; GENERAL INFORMATION:
; APPLICANT: Rosenfield, Robert L.
; TITLE OF INVENTION: IDENTIFICATION OF ACTIVATORS AND
; TITLE OF INVENTION: INHIBITORS OF SEBUM FORMATION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```


Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 CAGAAATGACCATGGTTGACACAG 42
|||||
Db 341 CTGAAATTACCATGGTTGACACAG 364

RESULT 11
US-08-465-375-1
; Sequence 1, Application US/08465375A
; Patent No. 6022897
; GENERAL INFORMATION:
; APPLICANT: Evans, Ronald M.
; APPLICANT: Forman, Barry M.
; TITLE OF INVENTION: SELECTIVE MODULATORS OF PEROXISOME
; TITLE OF INVENTION: PROLIFERATOR ACTIVATED RECEPTOR-GAMMA, AND METHODS FOR THE
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: SALK1470-1
; CURRENT APPLICATION NUMBER: US/08/465,375A
; CURRENT FILING DATE: 1995-06-05
; EARLIER APPLICATION NUMBER: 08/428,559
; EARLIER FILING DATE: 1995-04-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2005
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (352)...(1776)
US-08-465-375-1

Query Match 49.5%; Score 20.8; DB 3; Length 2005;
Best Local Similarity 91.7%; Pred. No. 11;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 CAGAAATGACCATGGTTGACACAG 42
|||||
Db 341 CTGAAATTACCATGGTTGACACAG 364

RESULT 12
US-09-788-070-1
; Sequence 1, Application US/09788070
; Patent No. 6413994
; GENERAL INFORMATION:
; APPLICANT: EVANS, Ronald
; APPLICANT: FORMAN, Barry
; TITLE OF INVENTION: MODULATORS OF PEROXISOME PROLIFERATOR ACTIVATED RECEPTOR-GAMMA,
; TITLE OF INVENTION: AND METHODS FOR THE USE THEREOF
; FILE REFERENCE: SALK1480-2
; CURRENT APPLICATION NUMBER: US/09/788,070
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/955,302
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 2005
; TYPE: DNA
; ORGANISM: Mus Musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (352)...(1776)
US-09-788-070-1

Query Match 49.5%; Score 20.8; DB 4; Length 2005;
Best Local Similarity 91.7%; Pred. No. 11;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 CAGAAATGACCATGGTTGACACAG 42
|||||

Db 341 CTGAAATTACCATGGTTGACACAG 364

RESULT 13
US-08-724-394A-20
; Sequence 20, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchinashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H. CONTIG"
US-08-724-394A-20

Query Match 47.1%; Score 19.8; DB 2; Length 246240;
Best Local Similarity 69.2%; Pred. No. 89;
Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 AGACCTACCCCGTGGGACAAATGACCATGGTTGACA 39
|||||
Db 35492 ATACTCACAGCGTCTCTGGAGAAATGACCATGTCCACA 35530

RESULT 14
US-08-724-394A-21
; Sequence 21, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 23:59:20 ; Search time 27 seconds
(without alignments)
693.573 Million cell updates/sec

Title: US-09-765-111A-7

Perfect score: 42

Sequence: 1 agacctaccctgggtggca.....aatgaccatggtgacacag 42

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/ECT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	42	100.0	42	10	US-09-765-111A-7
2	42	100.0	2334	10	US-09-765-111A-1
3	30.4	72.4	42	10	US-09-765-111A-11
4	30.4	72.4	2596	10	US-09-765-111A-22
5	30.4	72.4	2625	10	US-09-765-111A-5
6	30.2	71.9	1811	10	US-09-765-111A-26
7	30.2	71.9	1811	10	US-09-880-107-2212
8	30.2	71.9	2260	10	US-09-816-828-8
9	30.2	71.9	2329	10	US-09-816-828-9
10	25.2	60.0	42	10	US-09-765-111A-3
11	25.2	60.0	2523	10	US-09-765-111A-3
12	23	54.8	1608	10	US-09-765-111A-15
13	23	54.8	1679	12	US-10-109-886-5
14	22	52.4	343	9	US-10-046-935-990
15	22	52.4	343	9	US-09-878-178-990
16	22	52.4	611	10	US-09-925-299-109
17	22	52.4	683	10	US-09-867-701-6261
18	22	52.4	1372	10	US-09-765-111A-13
19	22	52.4	2711	10	US-09-765-111A-24

```

20 21.6 51.4 167 10 US-09-864-761-27653 Sequence 27653, A
21 21.6 51.4 455 10 US-09-864-761-11018 Sequence 11018, A
22 21.6 51.4 2951 12 US-10-044-090-587 Sequence 587, App
23 21.6 51.4 3305 12 US-10-044-090-586 Sequence 586, App
24 21 50.0 5385 10 US-09-920-804-1 Sequence 1, Appli
25 20.8 49.5 2005 12 US-10-142-373-1 Sequence 1, Appli
26 20.8 49.5 2793 9 US-09-338-842A-394 Sequence 394, App
27 20.6 49.0 364 10 US-09-864-761-18051 Sequence 18051, A
28 20.6 49.0 467 10 US-09-864-761-1290 Sequence 1290, Ap
29 20.6 49.0 1240 12 US-10-044-090-441 Sequence 441, App
30 20.2 48.1 1226 10 US-09-880-107-1503 Sequence 1503, Ap
31 20.2 48.1 1226 10 US-09-070-927A-711 Sequence 711, App
32 20.2 48.1 1590 9 US-09-738-626-688 Sequence 688, App
33 20.2 48.1 3164 10 US-09-764-869-1451 Sequence 1451, Ap
34 20.2 48.1 62944 10 US-09-954-456-2257 Sequence 2257, Ap
35 20 47.6 197 10 US-09-878-574-8482 Sequence 8482, Ap
36 20 47.6 268 10 US-09-294-093B-4731 Sequence 4731, Ap
37 20 47.6 506 10 US-09-833-381-1078 Sequence 1078, Ap
38 20 47.6 3453 9 US-10-108-605-262 Sequence 262, App
39 20 47.6 7305 10 US-09-795-693-9 Sequence 9, Appli
40 20 47.6 8056 9 US-10-072-621-3 Sequence 3, Appli
41 20 47.6 8195 10 US-09-795-693-7 Sequence 7, Appli
42 19.8 47.1 2364 9 US-09-712-363-32 Sequence 32, Appli
43 19.6 46.7 100 10 US-09-864-761-24461 Sequence 24461, A
44 19.6 46.7 289 10 US-09-864-761-17559 Sequence 17559, A
45 19.6 46.7 298 10 US-09-864-761-17170 Sequence 17170, A

```

ALIGNMENTS

RESULT 1

```

US-09-765-111A-7
; Sequence 7, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PPARGAMMA NUCLEIC ACID MOLECULES
; AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(41)
US-09-765-111A-7

```

Query Match 100.0%; Score 42; DB 10; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AGACCTACCCCGTGGTGACAGAAATGACCATGTTGACACAG 42
Db 1 AGACCTACCCCGTGGTGACAGAAATGACCATGTTGACACAG 42

```

RESULT 2

```

US-09-765-111A-1
; Sequence 1, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.

```

; APPLICANT: Kroll, Todd G.
 ; TITLE OF INVENTION: FAX8-PPARGAMMA NUCLEIC ACID MOLECULES
 ; FILE REFERENCE: B0801/7196/ERP/MAT
 ; CURRENT APPLICATION NUMBER: US/09/765,111A
 ; CURRENT FILING DATE: 2001-01-18
 ; PRIOR APPLICATION NUMBER: US 60/177,109
 ; PRIOR FILING DATE: 2000-01-20
 ; PRIOR APPLICATION NUMBER: US 60/225,079
 ; PRIOR FILING DATE: 2000-08-14
 ; NUMBER OF SEQ ID NOS: 47
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 2334
 ; TYPE: DNA
 ; ORGANISM: Homo Sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(2334)
 ; US-09-765-111A-1

Query Match 100.0%; Score 42; DB 10; Length 2334;
 Best Local Similarity 100.0%; Pred. No. 4.9e-08;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGACCTACCCGTCGTGGCAGAAATGACCATGTTGACACAG 42
 Db 878 AGACCTACCCGTCGTGGCAGAAATGACCATGTTGACACAG 919

RESULT 3

US-09-765-111A-11
 ; Sequence 11, Application US/09765111A
 ; Patent No. US20020106796A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fletcher, Jonathan A.
 ; APPLICANT: Kroll, Todd G.
 ; TITLE OF INVENTION: FAX8-PPARGAMMA NUCLEIC ACID MOLECULES
 ; FILE REFERENCE: B0801/7196/ERP/MAT
 ; CURRENT APPLICATION NUMBER: US/09/765, 11A
 ; CURRENT FILING DATE: 2001-01-18
 ; PRIOR APPLICATION NUMBER: US 60/177,109
 ; PRIOR FILING DATE: 2000-01-20
 ; PRIOR APPLICATION NUMBER: US 60/225,079
 ; PRIOR FILING DATE: 2000-08-14
 ; NUMBER OF SEQ ID NOS: 47
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 11
 ; LENGTH: 42
 ; TYPE: DNA
 ; ORGANISM: Homo Sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (3)...(41)
 ; US-09-765-111A-11

Query Match 72.4%; Score 30.4; DB 10; Length 42;
 Best Local Similarity 96.9%; Pred. No. 0.0011;
 Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 CTGTGTGGCAGAAATGACCATGTTGACACAG 42
 Db 11 CATGTGTGGCAGAAATGACCATGTTGACACAG 42

RESULT 4

US-09-765-111A-22
 ; Sequence 22, Application US/09765111A
 ; Patent No. US20020106796A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fletcher, Jonathan A.
 ; APPLICANT: Kroll, Todd G.

; TITLE OF INVENTION: FAX8-PPARGAMMA NUCLEIC ACID MOLECULES
 ; FILE REFERENCE: B0801/7196/ERP/MAT
 ; CURRENT APPLICATION NUMBER: US/09/765,111A
 ; CURRENT FILING DATE: 2001-01-18
 ; PRIOR APPLICATION NUMBER: US 60/177,109
 ; PRIOR FILING DATE: 2000-01-20
 ; PRIOR APPLICATION NUMBER: US 60/225,079
 ; PRIOR FILING DATE: 2000-08-14
 ; NUMBER OF SEQ ID NOS: 47
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 2596
 ; TYPE: DNA
 ; ORGANISM: Homo Sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (161)...(2596)
 ; US-09-765-111A-22

Query Match 72.4%; Score 30.4; DB 10; Length 2596;
 Best Local Similarity 96.9%; Pred. No. 0.0028;
 Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 CTGTGTGGCAGAAATGACCATGTTGACACAG 42
 Db 1150 CATGTGTGGCAGAAATGACCATGTTGACACAG 1181

RESULT 5

US-09-765-111A-5
 ; Sequence 5, Application US/09765111A
 ; Patent No. US20020106796A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fletcher, Jonathan A.
 ; APPLICANT: Kroll, Todd G.
 ; TITLE OF INVENTION: FAX8-PPARGAMMA NUCLEIC ACID MOLECULES
 ; FILE REFERENCE: B0801/7196/ERP/MAT
 ; CURRENT APPLICATION NUMBER: US/09/765,111A
 ; CURRENT FILING DATE: 2001-01-18
 ; PRIOR APPLICATION NUMBER: US 60/177,109
 ; PRIOR FILING DATE: 2000-01-20
 ; PRIOR APPLICATION NUMBER: US 60/225,079
 ; PRIOR FILING DATE: 2000-08-14
 ; NUMBER OF SEQ ID NOS: 47
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 2625
 ; TYPE: DNA
 ; ORGANISM: Homo Sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(2625)
 ; US-09-765-111A-5

Query Match 72.4%; Score 30.4; DB 10; Length 2625;
 Best Local Similarity 96.9%; Pred. No. 0.0028;
 Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 CTGTGTGGCAGAAATGACCATGTTGACACAG 42
 Db 1179 CATGTGTGGCAGAAATGACCATGTTGACACAG 1210

RESULT 6

US-09-765-111A-26
 ; Sequence 26, Application US/09765111A
 ; Patent No. US20020106796A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fletcher, Jonathan A.
 ; APPLICANT: Kroll, Todd G.
 ; TITLE OF INVENTION: FAX8-PPARGAMMA NUCLEIC ACID MOLECULES

; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF

; FILE REFERENCE: B0801/7196/ERP/MAT

; CURRENT APPLICATION NUMBER: US/09/765,111A

; CURRENT FILING DATE: 2001-01-18

; PRIOR APPLICATION NUMBER: US 60/177,109

; PRIOR FILING DATE: 2000-01-20

; PRIOR APPLICATION NUMBER: US 60/225,079

; PRIOR FILING DATE: 2000-08-14

; NUMBER OF SEQ ID NOS: 47

; SOFTWARE: Fast-Seq for Windows Version 3.0

; SEQ ID NO 26

; LENGTH: 1811

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (173)....(1609)

US 765-111A-26

Query Match 71.9%; Score 30.2; DB 10; Length 1811;

Best Local Similarity 91.4%; Pred. No. 0.0031;

Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCGTGTGGCAGAAATGACCATGGTTGACACAG 42

DB 157 CGCGTGTGGCAGAAATGACCATGGTTGACACAG 191

RESULT 7

US-09-880-107-2212

; Sequence 2212, Application US/09880107

; Patent No. US20020142981A1

; GENERAL INFORMATION:

; APPLICANT: Horne, Darci T.

; APPLICANT: Vockley, Joseph G.

; APPLICANT: Scherf, Uwe

; APPLICANT: Gene Logic, Inc.

; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

; FILE REFERENCE: 44921-5028-WO

; CURRENT APPLICATION NUMBER: US/09/880,107

; CURRENT FILING DATE: 2001-06-14

; PRIOR APPLICATION NUMBER: US 60/211,379

; PRIOR FILING DATE: 2000-06-14

; PRIOR APPLICATION NUMBER: US 60/237,054

; PRIOR FILING DATE: 2000-10-02

; NUMBER OF SEQ ID NOS: 3950

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2212

; LENGTH: 1811

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L40904

US-09-880-107-2212

Query Match 71.9%; Score 30.2; DB 10; Length 1811;

Best Local Similarity 91.4%; Pred. No. 0.0031;

Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCGTGTGGCAGAAATGACCATGGTTGACACAG 42

DB 157 CGCGTGTGGCAGAAATGACCATGGTTGACACAG 191

RESULT 8

US-09-816-828-8/c

; Sequence 8, Application US/09816828

; Patent No. US20020150898A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

US-09-816-828-8/c

Query Match 71.9%; Score 30.2; DB 10; Length 1811;

Best Local Similarity 91.4%; Pred. No. 0.0031;

Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCGTGTGGCAGAAATGACCATGGTTGACACAG 42

DB 157 CGCGTGTGGCAGAAATGACCATGGTTGACACAG 191

RESULT 9

US-09-816-828-8/c

; Sequence 9, Application US/09816828

; Patent No. US20020150898A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

US-09-816-828-8/c

Query Match 71.9%; Score 30.2; DB 10; Length 2260;

Best Local Similarity 91.4%; Pred. No. 0.0032;

Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCGTGTGGCAGAAATGACCATGGTTGACACAG 42

DB 1655 CGCGTGTGGCAGAAATGACCATGGTTGACACAG 1621

RESULT 9

US-09-816-828-8/c

; Sequence 9, Application US/09816828

; Patent No. US20020150898A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

US-09-816-828-8/c

Query Match 71.9%; Score 30.2; DB 10; Length 2260;

Best Local Similarity 91.4%; Pred. No. 0.0032;

Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCGTGTGGCAGAAATGACCATGGTTGACACAG 42

DB 1655 CGCGTGTGGCAGAAATGACCATGGTTGACACAG 1621

RESULT 9

US-09-816-828-8/c

; Sequence 9, Application US/09816828

; Patent No. US20020150898A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

US-09-816-828-8/c

Query Match 71.9%; Score 30.2; DB 10; Length 2260;

Best Local Similarity 91.4%; Pred. No. 0.0032;

Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCGTGTGGCAGAAATGACCATGGTTGACACAG 42

DB 1655 CGCGTGTGGCAGAAATGACCATGGTTGACACAG 1621

RESULT 9

US-09-816-828-8/c

; Sequence 9, Application US/09816828

; Patent No. US20020150898A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

US-09-816-828-8/c

Query Match 71.9%; Score 30.2; DB 10; Length 2260;

Best Local Similarity 91.4%; Pred. No. 0.0032;

Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCGTGTGGCAGAAATGACCATGGTTGACACAG 42

DB 1655 CGCGTGTGGCAGAAATGACCATGGTTGACACAG 1621

RESULT 9

US-09-816-828-8/c

; Sequence 9, Application US/09816828

; Patent No. US20020150898A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

US-09-816-828-8/c

Query Match 71.9%; Score 30.2; DB 10; Length 2260;

Best Local Similarity 91.4%; Pred. No. 0.0032;

Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCGTGTGGCAGAAATGACCATGGTTGACACAG 42

DB 1655 CGCGTGTGGCAGAAATGACCATGGTTGACACAG 1621

RESULT 9

US-09-816-828-8/c

; Sequence 9, Application US/09816828

; Patent No. US20020150898A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

US-09-816-828-8/c

Query Match 71.9%; Score 30.2; DB 10; Length 2260;

Best Local Similarity 91.4%; Pred. No. 0.0032;

Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCGTGTGGCAGAAATGACCATGGTTGACACAG 42

DB 1655 CGCGTGTGGCAGAAATGACCATGGTTGACACAG 1621

RESULT 9

US-09-816-828-8/c

; Sequence 9, Application US/09816828

; Patent No. US20020150898A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

US-09-816-828-8/c

Query Match 71.9%; Score 30.2; DB 10; Length 2260;

Best Local Similarity 91.4%; Pred. No. 0.0032;

Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCGTGTGGCAGAAATGACCATGGTTGACACAG 42

DB 1655 CGCGTGTGGCAGAAATGACCATGGTTGACACAG 1621

RESULT 9

US-09-816-828-8/c

; Sequence 9, Application US/09816828

; Patent No. US20020150898A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

US-09-816-828-8/c

Query Match 71.9%; Score 30.2; DB 10; Length 2260;

Best Local Similarity 91.4%; Pred. No. 0.0032;

Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCGTGTGGCAGAAATGACCATGGTTGACACAG 42

DB 1655 CGCGTGTGGCAGAAATGACCATGGTTGACACAG 1621

RESULT 9

US-09-816-828-8/c

; Sequence 9, Application US/09816828

; Patent No. US20020150898A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

US-09-816-828-8/c

Query Match 71.9%; Score 30.2; DB 10; Length 2260;

Best Local Similarity 91.4%; Pred. No. 0.0032;

Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCGTGTGGCAGAAATGACCATGGTTGACACAG 42

DB 1655 CGCGTGTGGCAGAAATGACCATGGTTGACACAG 1621

RESULT 9

US-09-816-828-8/c

; Sequence 9, Application US/09816828

; Patent No. US20020150898A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

US-09-816-828-8/c

Query Match 71.9%; Score 30.2; DB 10; Length 2260;

Best Local Similarity 91.4%; Pred. No. 0.0032;

Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCGTGTGGCAGAAATGACCATGGTTGACACAG 42

DB 1655 CGCGTGTGGCAGAAATGACCATGGTTGACACAG 1621

RESULT 9

US-09-816-828-8/c

; Sequence 9, Application US/09816828

; Patent No. US20020150898A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

US-09-816-828-8/c

Query Match 71.9%; Score 30.2; DB 10; Length 2260;

Best Local Similarity 91.4%; Pred. No. 0.0032;

Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCGTGTGGCAGAAATGACCATGGTTGACACAG 42

DB 1655 CGCGTGTGGCAGAAATGACCATGGTTGACACAG 1621

RESULT 9

US-09-816-828-8/c

; Sequence 9, Application US/09816828

; Patent No. US20020150898A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

US-09-816-828-8/c

Query Match 71.9%; Score 30.2; DB 10; Length 2260;

Best Local Similarity 91.4%; Pred. No. 0.0032;

Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Query Match 71.9%; Score 30.2; DB 10; Length 2329;
Best Local Similarity 91.4%; Pred. No. 0.0033;
Matches 32; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 8 CCCCCCGTGGCAGAAATGACCATGGTTGACACAG 42
DB 1655 CGCCGTGGCCGAGAAATGACCATGGTTGACACAG 1621

RESULT 10

US-09-765-111A-9
; Sequence 9, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PPARGAMMA NUCLEIC ACID MOLECULES
; AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/77,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(41)
US-09-765-111A-9

Query Match 60.0%; Score 25.2; DB 10; Length 42;
Best Local Similarity 79.9%; Pred. No. 0.15;
Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 CTACCCCGTGGCAGAAATGACCATGGTTGACACAG 42
DB 5 CCAGGCCCTCTCTCAGAAATGACCATGGTTGACACAG 42

RESULT 11

US-09-765-111A-3
; Sequence 3, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PPARGAMMA NUCLEIC ACID MOLECULES
; AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2523)
US-09-765-111A-3

Query Match 60.0%; Score 25.2; DB 10; Length 2523;
Best Local Similarity 79.9%; Pred. No. 0.37;
Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 CTACCCCGTGGCAGAAATGACCATGGTTGACACAG 42
DB 1071 CCAGGCCCTCTCTCAGAAATGACCATGGTTGACACAG 1108

RESULT 12

US-09-765-111A-15
; Sequence 15, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PPARGAMMA NUCLEIC ACID MOLECULES
; AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 1608
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (91)...(1608)
US-09-765-111A-15

Query Match 54.8%; Score 23; DB 10; Length 1608;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AGAATGACCATGGTTGACACAG 42
DB 171 AGAATGACCATGGTTGACACAG 193

RESULT 13

US-10-109-886-5
; Sequence 5, Application US/10109886
; Patent No. US20020119499A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO. LTD.
; APPLICANT: MIZUKAMI, Junko
; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND
; ANTAGONIST TO PPAR
; FILE REFERENCE: TANIUCHI-6
; CURRENT APPLICATION NUMBER: US/10/109,886
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/514,247
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/JP98/03734
; PRIOR FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: JP231084/1997
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)...(1679)

US-10-109-886-5

Query Match 54.8%; Score 23; DB 12; Length 1679;
Best Local Similarity 100.0%; Pred. No. 2.7;

20 AGAATGACCATGGTTGACACAG 42

RESULT 14
 S-10-046-935-990
 Sequence 990, Application US/10046935
 Patent No. US20020156011A1

Query Match 52.4%; Score 22; DB 9; Length 343;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 22; Conservative 0; Mismatches 0; Indels

21 GAAATGACCATGGTTGACACAG 42
121 GAAATGACCATGGTTGACACAG 142

RESULT 15
 87-178-990
 since 990, Application US/09878178
 Agent No. US20020177552A1
 GENERAL INFORMATION:
 APPLICANT: JIANG, YUQIU
 APPLICANT: Harlocker, Susan L.
 APPLICANT: Sectris, Heather
 TITLE OF INVENTION: COMPOSITIONS AND METHODS
 OF INVENTION: AND DIAGNOSIS OF COLON
 CANCER
 FILE REFERENCE: 210121.527
 CURRENT APPLICATION NUMBER: US/09/878,178
 CURRENT FILING DATE: 2001-06-08
 NUMBER OF SEQ ID NOS: 2237
 SOFTWARE: FastSeq For Windows Version 4.0
 SEQ ID NO 990
 LENGTH: 343

Query Match 52.4%; Score 22; DB 9; Length 343;
Best Local Similarity 100.0%; Pred.No. 4-8;
Matches 22; Conservative 0; Mismatches 0; Indels

21 GAAATGACCATGGTTGACACAG 42
121 GAAATGACCATGGTTGACACAG 142

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 23:02:45 ; Search time 1003.67 Seconds
(without alignments)
677.725 Million cell updates/sec

Title: US-09-765-111a-7
Perfect score: 42
Sequence: 1 agacctaccctggtggca.....aatgaccatggtgacacag 42

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Sequences: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_estc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gas.*
18: em_gas_hum.*
19: em_gas_inv.*
20: em_gas_pin.*
21: em_gas_vrt.*
22: em_gas_fun.*
23: em_gas_mam.*
24: em_gas_mus.*
25: em_gas_other.*
26: em_gas_pro.*
27: em_gas_rod.*

ALIGNMENTS

RESULT 1
AL543579
LOCUS
DEFINITION
AL543579 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1006Y110 5
prime, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
EST.
AL543579.1 GI:13876058
SOURCE
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Li W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source

Location/Qualifiers
1..903
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1006Y110"
/clone_lib="LTI_NFL006_PL2"
/cissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT)-primer. Five prime end

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30.2	71.9	903	9	AL543579
2	30.2	71.9	1141	14	BM922116
3	24	57.1	1243	14	BM924484
4	23	54.8	526	12	BF957583
5	23	54.8	901	13	BI820841
6	22.8	54.3	700	12	BF628555

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

HVCDNA0002 (Dehydration stress)"
 /tissue_type="Seedling shoot"
 /lab_host="TJCI21"
 /notes="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
 Seeds were surface sterilized then germinated under axenic
 conditions in the dark at room temperature on filter paper
 with water, nystatin and cerotaxime in covered
 crystallization dishes. Five-day old seedlings were
 incubated at 90% RH for 24 hr. Shoots were then harvested,
 total RNA was prepared, poly(A) RNA was purified, one
 primary unamplified cDNA library was made, 600000 pfu were
 in vivo excised to give pBluescript SK(-) cDNA phagemids.
 These steps were performed in the TU Close laboratory at
 the University of California, Riverside (Choi, Close,
 Fenton). Phagemids were plated and picked at the Clemson
 University Genomics Institute (CUGI) (Begum, Palmer,
 Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
 sequencing and sequence analysis were performed at CUGI
 (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
 The sequence has been trimmed to remove vector sequence
 and contains a minimum of 100 bases of phred value 20 or
 above. For more details on library preparation and
 sequence analysis see
<http://www.genome.clemson.edu/projects/barley>. To order
 this clone see <http://www.genome.clemson.edu/orders> Also
 see Close TJ, Wing R, Kleinhoft A, Wise R (2001)
 Genetically and physically anchored EST resources for
 barley genomics. Barley Genetics Newsletter 31:29-30.
 (<http://wheat.pw.usda.gov/ggpages/Lgn/31/cover.1.cml>)"

BASE COUNT 147 a 212 c 202 g 139 t
 ORIGIN

Query Match 54.3%; Score 22.8; DB 12; Length 700;
 Best Local Similarity 71.4%; Pred. No. 1.6e+02;
 Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1 AGACTACCCCGTGGCAGAAATGCACGTTGACACAG 42
 Db 457 ATACCTTCACTGCTGGCAAGAAACACCGTGTGACACAG 498

RESULT 7
 AW290590
 LOCUS
 DEFINITION NXNW031H10F Nsf Xylem Normal wood Vertical Pinus taeda cDNA clone
 ACCESSION AW290590
 VERSION AW290590.1 GI:6697226
 KEYWORDS EST.
 SOURCE loblolly pine.
 ORGANISM Pinus taeda

REFERENCE
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Coniferopsida; Coniferales; Pinus; Pinus;
 Sederoff, R.
 AUTHORS Molecular Basis of Wood Formation in the Pine Megagenome
 TITLE Unpublished (2000)
 JOURNAL North Carolina State University
 COMMENT Contact: Johnson, Arthur
 Tel: 919 515 7800
 Fax: 919 515 7801
 Email: ajohnson@unity.ncsu.edu
 Seq primer: T3.

FEATURES
 Location/Qualifiers
 1..463
 /organism="Pinus taeda"
 /db_xref="taxon:3352"
 /clone="NXNW031H10"

Qy 2 GACCTACCCCGTGGCAGAAATGCACGTTGACACAC 41
 Db 2 GGCATATCTCGAGTGGCAGTGAAGACCATGCTCACATA 41

RESULT 9
 AZ286381
 LOCUS
 DEFINITION RPCI-23-155J16.TJ RPCI-23 Mus musculus genomic clone RPCI-23-155J16

BASE COUNT 118 a 97 c 104 g 127 t 17 others
 ORIGIN

Query Match 53.3%; Score 22.4; DB 10; Length 463;
 Best Local Similarity 72.5%; Pred. No. 1.9e+02;
 Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 GACCTACCCCGTGGCAGAAATGCACGTTGACACAC 41
 Db 102 GGCATATCTCGAGTGGCAGTGAAGACCATGCTCACATA 141

RESULT 8
 BF609114
 LOCUS
 DEFINITION NXSI_040_E09 F NXSI (Nsf Xylem Side wood Inclined) Pinus taeda cDN
 clone NXSI_040_E09 5', mRNA sequence.
 ACCESSION BF609114
 VERSION BF609114.1 GI:11776445
 KEYWORDS EST.
 SOURCE loblolly pine.
 ORGANISM Pinus taeda

REFERENCE
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Coniferopsida; Coniferales; Pinus; Pinus;
 Sederoff, R.
 AUTHORS Molecular Basis of Wood Formation in the Pine Megagenome
 TITLE Unpublished (2000)
 JOURNAL Contact: Johnson, Arthur
 COMMENT North Carolina State University
 Tel: 919 515 7800
 Fax: 919 515 7801
 Email: ajohnson@unity.ncsu.edu
 Seq primer: T3.

FEATURES
 Location/Qualifiers
 1..514
 /organism="Pinus taeda"
 /strain="Coastal plain loblolly pine from North Carolina"
 /db_xref="taxon:3352"
 /clone="NXSI_040_E09"
 /clone_lib="NXSI (Nsf Xylem Side wood Inclined)"
 /tissue_type="Xylem"
 /cell_type="Side"
 /dev_stage="Juvenile"
 /lab_host="XLI-Blue"
 /notes="Vector: Bluescript SK; Site 1: Eco RI; Site 2: Xba
 ; The library is from early (spring) wood, taken from
 three six-year old trees (three different genotypes), in
 the juvenile phase. These trees were induced to form side
 wood by bending to a 45 degree angle and tying them to the
 ground. Differentiating xylem was harvested from the sides
 of the inclined stems, and a mixture of all three
 genotypes was used for the library. oligo-dr primed cDNA
 was directionally cloned into the EcoRI-XhoI Bluescript SK
 vector arms. NOTE: The sequences contain a 'cDNA adapter'
 between the EcoRI site and the start of the EST. The
 adapter sequence is 'AATTGGCACGAG'."

BASE COUNT 138 a 103 c 115 g 141 t 17 others
 ORIGIN

Query Match 53.3%; Score 22.4; DB 12; Length 514;
 Best Local Similarity 72.5%; Pred. No. 2e+02;
 Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 GACCTACCCCGTGGCAGAAATGCACGTTGACACAC 41
 Db 2 GGCATATCTCGAGTGGCAGTGAAGACCATGCTCACATA 41

RESULT 9
 AZ286381
 LOCUS
 DEFINITION RPCI-23-155J16.TJ RPCI-23 Mus musculus genomic clone RPCI-23-155J16

, DNA sequence.
 ACCESSION AZ286381
 VERSION AZ286381.1 GI:9528090
 KEYWORDS GSS
 SOURCE house mouse.
 ORGANISM Mus musculus
 REFERENCE 1 (bases 1 to 544)
 AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akineret,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Fraser,C.M.
 TITLE Mouse BAC End Sequences from Library RPCI-23
 JOURNAL Unpublished (1999)
 COMMENT Other GSSs: RPCI-23-155U16-TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 155 row: J column: 16
 Seq primer: SP6
 Class: BAC ends.
 BASE COUNT 135 a 108 c 158 g 143 t
 ORIGIN
 source
 Location/Qualifiers
 1..544
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-155U16"
 /clone_lib="RPCI-23"
 /sex="Female"
 /lab_host="DH10B"
 /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."
 Query Match 53.3%; Score 22.4; DB 17; Length 544;
 Best Local Similarity 81.2%; Pred. No. 2.le+02;
 Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 11 CTTGTGGCAGAAATGACCATGTTGACACAG 42
 Db 438 CTTGTGGAGGAATCACCATTGTTATACAG 469
 RESULT 10
 B92851/c
 LOCUS B92851 544 bp DNA linear GSS 25-JUN-1998
 DEFINITION CIT-HSP-2164L16.TR CIT-HSP Homo sapiens genomic clone 2164L16, DNA sequence.
 ACCESSION B92851
 VERSION B92851.1 GI:2975188
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 544)
 AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,

Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.
 Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)
 JOURNAL Unpublished (1998)
 COMMENT Other GSSs: CIT-HSP-2164L16.TF
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadam@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
 Seq primer: M13 Reverse
 Class: BAC ends.
 BASE COUNT 143 a 152 c 100 g 149 t
 ORIGIN
 source
 Location/Qualifiers
 1..544
 /organism="Homo sapiens"
 /db_xref="GDB:7100882"
 /db_xref="taxon:9606"
 /clone="2164L16"
 /clone_lib="CIT-HSP"
 /sex="Male"
 /cell_type="Sperm"
 /note="vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"
 Query Match 53.3%; Score 22.4; DB 17; Length 544;
 Best Local Similarity 81.2%; Pred. No. 2.le+02;
 Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 8 CCCCCTGGTGGCAGAAATGACCATGTTGACA 39
 Db 386 CTTCTCTGGGGCAGAAATGCCCTCGTTGACA 355
 RESULT 11
 BQ655082
 LOCUS BQ655082 648 bp mRNA linear EST 15-JUL-2002
 DEFINITION NXSV090.A07.F NXRV (Nsf Xylem Root wood vertical) pinus taeda cDNA clone NXSV090.A07.5', mRNA sequence.
 ACCESSION BQ655082
 VERSION BQ655082.1 GI:21787408
 KEYWORDS EST.
 SOURCE loblolly pine.
 ORGANISM Pinus taeda
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
 REFERENCE 1 (bases 1 to 648)
 AUTHORS Sederoff,R.
 TITLE Molecular Basis of Wood Formation in the Pine Megagenome
 JOURNAL Unpublished (2000)
 COMMENT Contact: Johnson, Arthur
 North Carolina State University
 Tel: 919 515 7800
 Fax: 919 515 7801
 Email: ajohnson@unity.ncsu.edu
 Seq primer: T3
 BASE COUNT 143 a 152 c 100 g 149 t
 ORIGIN
 source
 Location/Qualifiers
 1..648
 /organism="Pinus taeda"
 /strain="Coastal plain loblolly pine from North Carolina"
 /db_xref="taxon:3352"
 /clone="NXSV090.A07"
 /clone_lib="NXRV (Nsf Xylem Root wood vertical)"
 /tissue_type="Xylem"
 /cell_type="Root (primary)"
 /dev_stage="Transitional"

/lab_host="XLI-Blue"
 /note="Vector: pBluescript SK-; Site 1: Eco RI; Site 2:
 xhoI; The library is from primary xylem scraped from the
 roots of a twelve year old tree in the transitional phase
 from juvenile wood to mature wood production. NOTE: The
 sequences contain a 'cDNA adapter' between the EcoRI site
 and the start of the EST. The adapter sequence is
 'AATTCGGCAGCAG'."

BASE COUNT 166 a 127 c 152 g 182 t 21 others
 ORIGIN
 Query Match 53.3%; Score 22.4; DB 14; Length 648;
 Best Local Similarity 72.5%; Pred. No. 2.3e+02;
 Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GACCTACCCCGTGGCAGAAATGACCATGGTGCACACA 41
 Db 29 GGCATATCCTGAGTGCCAGTGAAGACCATGGCTCACATA 68
 RESULT 12
 Locus CNS06RHC 991 bp DNA linear GSS 05-JUL-2001
 DEFINITION T3 end of clone AW0AA014D09 of library AW0AA from strain CLIB 89 of
 Yarrowia lipolytica, genomic survey sequence.
 ACCESSION AL411958
 VERSION AL411958.1 GI:12181936
 KEYWORDS GSS.
 SOURCE Yarrowia lipolytica.
 ORGANISM Yarrowia lipolytica.
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Dipodascaceae; Yarrowia.

REFERENCE 1 (bases 1 to 991)
 AUTHORS Souciet J.L., Aigle M., Artiguenave F., Blandin G.,
 Bolotin-Fukuhara M., Bon E., Broutier P., Casaregola S.,
 de-Montigny J., Dujon B., Durrens P., Lepingle A., Llorente B.,
 Malpertuy A., Neuveglise C., Ozier-Kalogeropoulos O., Potier S.,
 Saurin W., Tekai F., Toffano-Nioche C., Wesolowski-Louvel M.,
 Winkler P. and Weissenbach J.
 TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
 yeast species for molecular evolution studies
 JOURNAL FEMS Lett. 487 (1), 3-12 (2000)
 MEDLINE 20584711
 PUBMED 11152876
 REFERENCE 2 (bases 1 to 991)
 AUTHORS Casaregola S., Neuveglise C., Lepingle A., Bon E., Feynerol C.,
 Artiguenave F., Winkler P. and Gaillardin C.
 TITLE Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia
 lipolytica

JOURNAL FEMS Lett. 487 (1), 95-100 (2000)
 MEDLINE 20584727
 PUBMED 11152892
 REFERENCE 3 (bases 1 to 991)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
 2 rue Gaston Creneau, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
 seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 COMMENT This GSS is part of a random genomic sequencing program of thirteen
 yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
 exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
 Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
 lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
 angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
 Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
 5 kb were prepared and both extremities were sequenced. See
 keywords for description of this sequence and for the sequence of
 the other extremity of this insert.

FEATURES
 source
 1. .991
 Location/Qualifiers
 /organism="Yarrowia lipolytica"
 /strain="CLIB 89"
 /db_xref="taxon:4952"

/clone="AW0AA014D09"
 /clone_lib="AW0AA"
 /note="end : 73"
 BASE COUNT 257 a 198 c 258 g 275 t 3 others
 ORIGIN

Query Match 53.3%; Score 22.4; DB 17; Length 991;
 Best Local Similarity 72.5%; Pred. No. 2.9e+02;
 Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 3 ACTTACCCCGTGGCAGAAATGACCATGGTGCACACAG 42
 Db 764 ACTAAACCTGTAGTGGCAGAGTTACCATGGCTGTAAC 803

RESULT 13
 Locus BF652561 490 bp mRNA linear EST 25-APR-200.
 DEFINITION 276153 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION BF652561
 VERSION BF652561.1 GI:11917693
 KEYWORDS EST.
 SOURCE COW.
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Butheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 490)
 AUTHORS Smith T.P.L., Grosse W.M., Freking B.A., Roberts A.J., Stone R.T.,
 Casas E., Wray J.B., White J., Cho J., Fahrenkrug S.C., Bennett
 G.L., Heaton M.P., Laegreid W.W., Rohrer G.A., Chitko-McKown C.G.,
 Petrea G., Holt I., Karamycheva S., Liang F., Quackenbush J. and
 Keefe J.W.
 TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
 JOURNAL Genome Res. 11 (4), 626-630 (2001)
 MEDLINE 21180013
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore
 and -minmatch 12 options.
 PCR Primers
 FORWARD: AGGAACACGCTATGACCAT
 BACKWARD: GTTTCAGTCAGCAGC
 Plate: 66 row: K column: 2
 Seq primer: APTAGTGACACTATAG.

FEATURES
 source
 1. .490
 Location/Qualifiers
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_lib="MARC 3BOV"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pGVV SPORT6; Site 1: NotI; Site 2: SalI;
 Library made from pooled tissue from marrow, alveolar
 macrophage, ovary, fetal semitendinosus muscle, and fetal
 longissimus muscle."
 BASE COUNT 96 a 168 c 142 g 84 t
 ORIGIN

Query Match 52.9%; Score 22.2; DB 12; Length 490;
 Best Local Similarity 88.9%; Pred. No. 2.3e+02;
 Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AGACCTACCCCGTGGCAGAAATGA 27
 Db 114 AGACCTACCCCGTGGCAGGTACGA 140

```

RESULT 14
BJ493993
LOCUS      558 bp      mRNA      linear      EST 08-AUG-2002
DEFINITION BJ493993 MF01FSA cDNA Oryzias latipes cDNA clone MF01FSA035D13 5',
            mRNA sequence.
ACCESSION  BJ493993
VERSION     BJ493993.1 GI:22145919
KEYWORDS   EST.
SOURCE     Japanese medaka.
ORGANISM   Oryzias latipes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Notoleosteoi;
            Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
            Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE  1 (bases 1 to 558)
            Kohara,Y., Shin-I,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
            Medaka EST Project in Takeda's lab
            Unpublished (2001)
            Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source           1..558
                     /organism="Oryzias latipes"
                     /strain="d-rR"
                     /db_xref="taxon:8090"
                     /clone="MF01FSA035D13"
                     /clone_lib="MF01FSA cDNA"
                     /sex="mixture of female and male"
                     /tissue_type="whole embryo"
                     /dev_stage="fry stage 40"
BASE COUNT  158 a 114 c 114 g 171 t 1 others
ORIGIN
1..558
/organism="Oryzias latipes"
/strain="d-rR"
/db_xref="taxon:8090"
/clone="MF01FSA035D13"
/clone_lib="MF01FSA cDNA"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fry stage 40"
Query Match      52.9%; Score 22.2; DB 13; Length 558;
Best Local Similarity 77.1%; Pred. No. 2.5e+02;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY  8  CCCCTGCTGGCGAGAAATGACCATGGTTGACACAG 42
    |||||
DI  456 CCACGGGGGGCAATAGCGACCATGATTGACACAG 490

RESULT 15
AZ084869
LOCUS      569 bp      DNA      linear      GSS 08-MAY-2000
DEFINITION RPCI-23-35K15-TV RPCI-23 Mus musculus genomic clone RPCI-23-35K15,
            DNA sequence.
ACCESSION  AZ084869
VERSION     AZ084869.1 GI:7726602
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 569)
            Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akirret
            , B., Levins,M., McGann,G., Taegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
            Mouse BAC End Sequences from Library RPCI-23
            Unpublished (1999)
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org

```

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@jong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
 Plate: 35 row: X column: 15
 Seq primer: T7
 Class: BAC ends.

FEATURES

```

            Location/Qualifiers
     source           1..569
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-35K15"
                     /clone_lib="RPCI-23"
                     /sex="Female"
                     /lab_host="DH10B"
                     /note="Organ: Kidney/Brain; Vector: pBAC3.6; Site 1:
                     EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and partially digested
                     with a combination of EcoRI and EcoRI Methylase. Size
                     selected DNA was cloned into the pBAC3.6 vector at the
                     EcoRI sites. The ligation products were transformed into
                     DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT      163 a 120 c 143 g 143 t
ORIGIN

```

```

Query Match      52.9%; Score 22.2; DB 17; Length 569;
Best Local Similarity 77.1%; Pred. No. 2.5e+02;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY  2  GACCTACCCCGTGGTGGCAGAAATGACCATGGTTG 36
    |||||
Db  152 GACCTTCACGGGGTGGTGGAAATGACTTGGGTG 186

```

Search completed: January 22, 2003, 00:39:33
 Job time : 1006.67 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 23:02:20 ; Search time 775.333 Seconds
(without alignments)
1576.506 Million cell updates/sec

Title: US-09-765-111a-9

Perfect score: 42

Sequence: 1 cggcgccaggccctctctca.....aatgaccatggtgacacag 42

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

T: number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_em.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	28.4	67.6	416	9	AB005521S1	AB005521 Homo sapi
2	28.4	67.6	135667	9	AC093174	AC093174 Homo sapi
C 3	28.4	67.6	166043	9	AC090947	AC090947 Homo sapi
C 4	28.4	67.6	185608	9	AC027126	AC027126 Homo sapi
5	26.6	63.3	172445	9	AC009471	AC009471 Homo sapi
6	26	61.9	1766	9	HSPARGAM	X90563 H.sapiens m
7	26	61.9	1808	9	HUMPPARG	L40304 Homo sapien
8	26	61.9	1811	6	AX409565	AX409565 Sequence
9	26	61.9	1844	6	ARI39020	ARI39020 Sequence
10	24.2	57.6	2028	9	AY048697	AY048697 Macaca fa
11	24.2	57.6	2063	9	AY048699	AY048699 Macaca fa
12	23.8	56.7	135011	2	AC103039	AC103039 Rattus no
13	23.8	56.7	182241	2	AC112596	AC112596 Rattus no
C 14	23.6	56.2	206156	2	AC004387	AC004387 Homo sapi
C 15	23.6	56.2	209317	9	AL672032	AL672032 Human DNA
16	23.4	55.7	1711	4	BTPPARG1	Y12419 B.taurus mr
17	23.4	55.7	4065	12	AY071819	AY071819 Synthetic
18	23.4	55.7	173755	2	AC121976	AC121976 Mus muscu
C 19	23.4	55.7	173935	2	AC027797	AC027797 Homo sapi
C 20	23.4	55.7	187088	2	AC087737	AC087737 Homo sapi
C 21	23.4	55.7	188439	9	AC025919	AC025919 Homo sapi
C 22	23.4	55.7	221647	10	AL591003	AL591003 Mouse DNA
23	23.4	55.7	237588	2	AC025581	AC025581 Mus muscu
24	23.2	55.2	1242	4	CCPAX8G	X83592 C.canis Pax
25	23.2	55.2	1361	10	MPAX89	X98596 M.musculus
26	23.2	55.2	1380	4	CCPAX8A	X83591 C.canis Pax
27	23.2	55.2	179150	2	AC099634	AC099634 Mus muscu
C 28	23.2	55.2	179937	9	AC016683	AC016683 Homo sapi
C 29	23.2	55.2	235312	2	AL732528	AL732528 Mus muscu
30	23	54.8	1518	9	HSU63415	U63415 Human perox
31	23	54.8	1608	6	ARI21467	ARI21467 Sequence
32	23	54.8	1608	9	HSU79012	U79012 Human ligan
33	23	54.8	1679	6	AR203332	AR203332 Sequence
34	23	54.8	1679	9	HUMPPARG	D83233 Homo sapien
35	23	54.8	104314	9	AL512883	AL512883 Human DNA
36	23	54.8	106759	2	AC099083	AC099083 Rattus no
37	23	54.8	189763	2	AL589992	AL589992 Homo sapi
C 38	23	54.8	202822	2	AC034116	AC034116 Mus muscu
39	22.8	54.3	73615	2	AC025356	AC025356 Homo sapi
C 40	22.8	54.3	117026	9	HS329A5	Z97832 Human DNA s
C 41	22.8	54.3	129155	9	AL139244	AL139244 Human DNA
42	22.8	54.3	143749	9	AC004962	AC004962 Homo sapi
43	22.8	54.3	208571	9	AC010906	AC010906 Homo sapi
44	22.8	54.3	221640	2	AC123971	AC123971 Lemur cat
45	22.6	53.8	1128	10	AF345993	AF345993 Rattus no

ALIGNMENTS

RESULT 1
AB005521S1
LOCUS
DEFINITION
Homo sapiens ppar gamma gene for peroxisome proliferator
activated-receptor gamma, exon 1.
ACCESSION
AB005521
VERSION
AB005521.1 GI:2605490
KEYWORDS
ppar gamma; ppar gamma common exon1; peroxisome proliferator
activated-receptor gamma.
SEGMENT
1 of 6
SOURCE
Homo sapiens placenta DNA.
ORGANISM
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1

AUTHORS Okazawa,H., Mori,H., Tamori,Y., Araki,S., Niki,T., Masugi,J., Kawanishi,M., Kubota,T., Siroda,H. and Kasuga,M.
TITLE No coding mutations are detected in the peroxisome proliferator-activated receptor- gene in Japanese patients with lipotrophic diabetes
JOURNAL Diabetes (1997) In press
REFERENCE 2 (bases 1 to 416)
AUTHORS Okazawa,H.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-1997) Hideki Okazawa, Kobe University School of Medicine, 2nd Department of Internal Medicine; 7-5-1 Kusunoki-cho Chuo-Ku, Kobe 650, Japan (E-mail:okazawamed.kobe-u.ac.jp, Tel:81-78-341-7451, Fax:81-78-382-2080)

FEATURES Location/Qualifiers
 1. 416
 /organism="Homo sapiens"
 /db xref="taxon:9606"
 /tissue_type="placenta"
 <115..340
 /gene="ppar gamma"
 /product="peroxisome proliferator activated-receptor gamma"
 /notes="ppar gamma common exon1"
 /number=1

BASE COUNT 113 a 105 c 67 g 131 t
ORIGIN

Query Match 67.6%; Score 28.4; DB 9; Length 416;
 Best Local Similarity 96.7%; Pred. No. 1.3;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 TCCTCTCAGAAATGACCATGGTTGACACAG 42
 |||||
 DB 104 TCCTTTCAGAAATGACCATGGTTGACACAG 133

RESULT 2
AC093174 135667 bp DNA linear PRI 13-AUG-2001
LOCUS Homo sapiens chromosome 3 clone RP11-167M22 map 3p, complete
DEFINITION sequence.
AC093174 GI:15148930
VERSION AC093174.1
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 135667)
 Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D., Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,F., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L., Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J., Wang,J., Wang,J., Wang,L., Wang,R., Wang,R., Wang,R., Wang,X., Wang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N., Yu,J. and Yang,H.

TITLE Chromosome 3p genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 135667)
AUTHORS Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D., Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,F., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L., Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,R., Wang,R., Wang,R., Wang,X., Wang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N., Yu,J. and Yang,H.

TITLE Yu,J. and Yang,H.
JOURNAL Direct Submission
COMMENT Submitted (13-AUG-2001) Human Genomic Center, Institute of Genetics, Chinese Academy of Sciences, Datun Road, Beijing, 100101, P.R.China
 -----Genome Center
 Center:Beijing Center
 Center code:Beijing
 Website:http://hgsc.igtp.ac.cn
 http://www.genomics.org.cn
 Contact:hgsc@igtp.ac.cn
 ----- Project Information
 Center project name:13 project
 Center clone name: RP11-167M22
 ----- Summary Statistics
 Sequencing vector: pUC18; 100% of reads
 Chemistry: Dye-terminator; ET 55% of reads
 Chemistry: Dye-terminator Big Dye; 45% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 586 bases at least Q40
 Consensus quality: 919 bases at least Q30
 Consensus quality: 1154 bases at least Q20
 Insert size: 1198; sum-of-contigs
 Quality coverage: 1.48x in Q20 bases;sum-of-contigs

FEATURES Location/Qualifiers
 1. 135667
 /organism="Homo sapiens"
 /db xref="taxon:9606"
 /chromosome="3"
 /map="3p"
 /clone="RP11-167M22"
 41244 a 25375 c 26554 g 42494 t

BASE COUNT 41244 a 25375 c 26554 g 42494 t
ORIGIN

Query Match 67.6%; Score 28.4; DB 9; Length 135667;
 Best Local Similarity 96.7%; Pred. No. 0.76;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 TCCTCTCAGAAATGACCATGGTTGACACAG 42
 |||||
 DB 109918 TCCTTTCAGAAATGACCATGGTTGACACAG 109947

RESULT 3
AC090947 166043 bp DNA linear PRI 20-MAR-2001
LOCUS Homo sapiens chromosome 3 clone RP11-30G23 map 3p, complete
DEFINITION sequence.
AC090947 AC016333
VERSION AC090947.1 GI:113384351
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 166043)
 Wu,Q., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D., Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,F., Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y., Luo,J., Niu,Y., Qi,Q., Qi,X., Song,L., Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J., Wang,J., Wang,J., Wang,L., Wang,R., Wang,R., Wang,R., Wang,X., Wang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N., Yu,J. and Yang,H.

TITLE Chromosome 3p genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 166043)
AUTHORS Wu,Q., Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H., Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D.,

Guo,Z., He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C.,
 Li,F., Li,G., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B.,
 Liu,Y., Li,W., Li,X., Li,Y., Luo,J., Niu,Y., Qi,O., Qi,X., Song,L.,
 Song,S., Sun,M., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
 Wang,J., Wang,J., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X.,
 Wang,Y., Wu,D., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X., Yu,B.,
 Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
 Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
 Yu,J. and Yang,H.

Direct Submission
 Submitted (20-MAR-2001) Human Genomic Center, Institute of
 Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
 100101, P.R.China

On Mar 20, 2001 this sequence version replaced gi:8072582.

-----Genome Center
 Center:Beijing Center
 Center code:Beijing
 Website:http://hgsc.igtp.ac.cn
 http://www.genomics.org.cn
 Contact:hgc@igtp.ac.cn

-----Project Information
 Center project name:1% project
 Center clone name: RP11-30G23

-----Summary Statistics
 Sequencing vector: pUC18; 100% of reads
 Chemistry: Dye-terminator; ET 55% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 499 bases at least Q40
 Consensus quality: 605 bases at least Q30
 Consensus quality: 674 bases at least Q20
 Insert size: 692; sum-of-contigs
 Quality coverage: 2.80x in Q20 bases; sum-of-contigs

TITLE
 JOURNAL

COMMENT

FEATURES
 Source

BASE COUNT 48570 a 35113 c 34446 g 47914 t

Query Match 67.6%; Score 28.4; DB 9; Length 166043;
 Best Local Similarity 96.7%; Pred. No. 0.74;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Q. 13 TCCTCTCAGAAATGACCATGGTTGACACAG 42

Db 117082 TCCTTTGAGAAATGACCATGGTTGACACAG 117053

RESULT 4
 AC027126/c 185608 bp DNA linear PRI 06-MAR-2001
 LOCUS
 DEFINITION Homo sapiens chromosome 3 clone RP11-586C12 map 3p, complete
 sequence.
 AC027126
 VERSION AC027126.4 GI:13236635
 KEYWORDS HTG.
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 185608)
 Authors Wang, X., Zhang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X.,
 Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M.,
 Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,
 Yu, J. and Yang, H.

Wang, X., Zhang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X.,
 Yu, B., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M.,
 Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,
 Yu, J. and Yang, H.

Chromosome 3p genomic sequence
 Unpublished
 2 (bases 1 to 185608)
 Submitted (28-MAR-2000) Human Genomic Center, Institute of
 Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
 100101, P.R.China

-----Genome Center
 Center:Beijing Center
 Center code:Beijing
 Website:http://hgsc.igtp.ac.cn
 http://www.genomics.org.cn
 Contact:hgc@igtp.ac.cn

-----Project Information
 Center project name:1% project
 Center clone name: RP11-586C12

-----Summary Statistics
 Sequencing vector: pUC18; 100% of reads
 Chemistry: Dye-terminator; ET 55% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 187858 bases at least Q40
 Consensus quality: 189056 bases at least Q30
 Consensus quality: 189280 bases at least Q20
 Insert size: 185608; sum-of-contigs
 Quality coverage: 10.96x in Q20 bases; sum-of-contigs

TITLE
 JOURNAL

REFERENCE
 AUTHORS

Tao, R., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B.,
 Bao, X., Sun, Y., Wu, O., Wang, H., Yang, X., Cheng, C., Wang, Y., Niu, Y.,
 Qi, X., Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y.,
 Li, G., Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D.,
 Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J.
 and Yang, H.

Direct Submission
 Submitted (28-MAR-2000) Human Genomic Center, Institute of
 Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
 100101, P.R.China

-----Genome Center
 Center:Beijing Center
 Center code:Beijing
 Website:http://hgsc.igtp.ac.cn
 http://www.genomics.org.cn
 Contact:hgc@igtp.ac.cn

-----Project Information
 Center project name:1% project
 Center clone name: RP11-586C12

-----Summary Statistics
 Sequencing vector: pUC18; 100% of reads
 Chemistry: Dye-terminator; ET 55% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 187858 bases at least Q40
 Consensus quality: 189056 bases at least Q30
 Consensus quality: 189280 bases at least Q20
 Insert size: 185608; sum-of-contigs
 Quality coverage: 10.96x in Q20 bases; sum-of-contigs

TITLE
 JOURNAL

COMMENT

On Mar 6, 2001 this sequence version replaced gi:8101246.

-----Genome Center
 Center:Beijing Center
 Center code:Beijing
 Website:http://hgsc.igtp.ac.cn
 http://www.genomics.org.cn
 Contact:hgc@igtp.ac.cn

-----Project Information
 Center project name:1% project
 Center clone name: RP11-586C12

-----Summary Statistics
 Sequencing vector: pUC18; 100% of reads
 Chemistry: Dye-terminator; ET 55% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 187858 bases at least Q40
 Consensus quality: 189056 bases at least Q30
 Consensus quality: 189280 bases at least Q20
 Insert size: 185608; sum-of-contigs
 Quality coverage: 10.96x in Q20 bases; sum-of-contigs

FEATURES
 Source

BASE COUNT 53641 a 40367 c 39222 g 52378 t

ORIGIN

Query Match 67.6%; Score 28.4; DB 9; Length 185608;
 Best Local Similarity 96.7%; Pred. No. 0.74;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 TCCTCTCAGAAATGACCATGGTTGACACAG 42

Db 161299 TCCTTTGAGAAATGACCATGGTTGACACAG 161270

RESULT 5

AC009471 172445 bp DNA linear PRI 07-NOV-2001
 LOCUS Homo sapiens BAC clone RP11-168K7 from 2, complete sequence.
 DEFINITION AC009471
 ACCESSION AC009471
 VERSION AC009471.5 GI:13435274
 HTG.
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 172445)
 Sulston, J.E. and Waterston, R.

Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)

99063792
 PUBMED 9847074

2 (bases 1 to 172445)

Scott, K., Kyung, K., Maupin, R., Laplant, Y. and Reitz, L.

The sequence of Homo sapiens BAC clone RP11-168K7

Unpublished

3 (bases 1 to 172445)

Waterston, R.H.

Direct Submission

Submitted (24-AUG-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

4 (bases 1 to 172445)

Waterston, R.H.

Direct Submission

Submitted (23-MAR-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

5 (bases 1 to 172445)

Waterston, R.

Direct Submission

Submitted (09-AUG-2001) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

6 (bases 1 to 172445)

Waterston, R.

Direct Submission

Submitted (07-NOV-2001) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Mar 23, 2001 this sequence version replaced gi:9838098.

COMMENT

----- Genome Center

Center: Washington University Genome Sequencing Center

Web site: <http://genome.wustl.edu/gsc>

Contact: sapiens@watson.wustl.edu

----- Summary Statistics

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D.
 McPherson, Department of Genetics, Washington University, St. Louis
 MO. For additional information about the map position of this
 sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male
 donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E.,
 Tateo, M., Catanese, J.J. and de Jong, P.J. (1998) An improved
 approach for construction of bacterial artificial chromosome
 libraries. Genomics 51:1-8. The clone may be obtained either from
 Research Genetics, Inc. (<http://www.resgen.com>) or Peter de Jong
 and coworkers at the Roswell Park Cancer Institute
 (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

Actual start of this clone is at base position 1 of RP11-168K7;
 actual end is at base position 172445 of RP11-168K7.

The sequence RP11-168K7 from base position 72908 to 74615 cannot be
 guaranteed due to a tandem repeat. The tandem size is believed to
 be in agreement with digest information.

The sequence RP11-168K7 contains a dinucleotide (CT) repeat from
 base position 120454 to 120643 for which the exact length is
 unknown. Assembly of the database is consistent with digest
 information.

FEATURES

Source

1. .172445
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2"

/clone="RP11-168K7"

/clone_lib="RP11-11"

1. .858

/rpt_family="MER2_type"

repeat_region

1813. .2157

misc_feature

1821. .2411

misc_feature

1821. .2209

misc_feature

1821. .2202

misc_feature

1824. .2385

misc_feature

1828. .2565

misc_feature

1828. .2187

misc_feature

1831. .2002

misc_feature

1831. .1990

misc_feature

1832. .2299

misc_feature

1833. .2209

misc_feature

2000. .2350

misc_feature

2049. .2506

misc_feature

2069. .2674

misc_feature

2122. .2436

misc_feature

2155. .2444

misc_feature

2311. .2481

misc_feature

2314. .2549

misc_feature

2346. .2836

misc_feature

2418. .3064

misc_feature

2423. .2863

misc_feature

/note="similar to

EST H82730 (NID:g1060819) yv81d08.s1"

/note="similar to

EST A1819855 (NID:g5438934) wj45g03.x1"

/note="similar to

EST AA865426 (NID:g2957702) ch29e02.s1"

/note="similar to

EST A1335633 (NID:g4072560) qt24f02.x1"

/note="similar to

EST T59958 (NID:g661795) yc01h05.s1"

/note="similar to

EST BF342851 (NID:g11289878) "

/note="similar to

EST A1347836 (NID:g4085042) qp03a08.x1"

/note="similar to

EST A1266142 (NID:g3884300) qp64c04.x1"

/note="similar to

EST C21474 (NID:g1622584) "

/note="similar to

EST AA625085 (NID:g2537470) af67a07.r1"

/note="similar to

EST A1989581 (NID:g5836462) ws34f08.x1"

/note="similar to

EST T79044 (NID:g697553) YG17f05.r1"

/note="similar to

EST W76395 (NID:g1386619) zd66c01.r1"

/note="similar to

EST AW886527 (NID:g8048539) "

/note="similar to

EST BF845185 (NID:g12201486) "

/note="similar to

EST H83604 (NID:g1062275) yv81d08.r1"

/note="similar to

EST AW798977 (NID:g7850847) "

/note="similar to

EST N36581 (NID:g1157723) YX86e05.r1"

/note="similar to

EST AA878397 (NID:g2987362) oe65d06.s1"

/note="similar to

EST BF204462 (NID:g11098048) "

/note="similar to

EST AA128499 (NID:g1688496) z124c06.r1"

misc_feature 3602..2656 EST AI347836 (NID:g4085042) gp03a08.x1
/note="similar to
misc_feature 2676..3146 EST AI188866 (NID:g1775893) zp7e05.r1
/note="similar to
misc_feature 2685..2741 EST AI134393 (NID:g10994932)
/note="similar to
misc_feature 2704..3039 EST H80249 (NID:g1058338) yu87g12.r1
/note="similar to
misc_feature 2763..2832 EST AI335633 (NID:g4072560) qz24f02.x1
/note="similar to
repeat_region 2812..2849
/rptc_family="L2"
misc_feature 2816..3061 EST AA910154 (NID:g3049444) od92h01.s1
/note="similar to
misc_feature 3101..3634 EST AI754458 (NID:g5132722) cr25e10.x1
/note="similar to
misc_feature 3120..3839 EST AI186513 (NID:g3737151) qd35a07.x1
/note="similar to
misc_feature 3120..3708 EST AA639852 (NID:g2563631) ng85c10.s1
/note="similar to
misc_feature 3121..3272 EST AA723217 (NID:g2740854) zg80d09.s1
/note="similar to
misc_feature 3122..3760 EST AA196274 (NID:g1791838) zp92a03.s1
/note="similar to
misc_feature 3122..3650 EST AA716174 (NID:g2728448) zg93g05.s1
/note="similar to
misc_feature 3122..3527 EST AU153039 (NID:g1014560)
/note="similar to
misc_feature 3122..3506 EST T54461 (NID:g656322) yb06b02.s1
/note="similar to
misc_feature 3140..3594 EST AA450105 (NID:g2163855) zx42d07.s1
/note="similar to
misc_feature 3171..3249 EST AA242948 (NID:g1873743) zt65b10.r1
/note="similar to
misc_feature 3191..3712 EST AV656934 (NID:g9877948)
/note="similar to
misc_feature 3193..4031 EST BE834287 (NID:g10266769)
/note="similar to
misc_feature 3290..3560 EST BE748820 (NID:g10162812)
/note="similar to
misc_feature 3306..3564 EST AI799528 (NID:g5365000) tr49b11.x1
/note="similar to
misc_feature 3312..3844 EST AV656934 (NID:g9877948)
/note="similar to
misc_feature 3331..3938 EST AA669825 (NID:g2631324) ag36e11.s1
/note="similar to
misc_feature 3331..3934 EST AW377487 (NID:g6882150)
/note="similar to
misc_feature 3340..4000 EST AW852068 (NID:g7947585)
/note="similar to
misc_feature 3355..3618 EST AW852329 (NID:g7947846)
/note="similar to
misc_feature 3359..3695 EST AA495727 (NID:g2229048) zw04a08.r1
/note="similar to
misc_feature 3539..3695 EST BE569839 (NID:g10082999)
/note="similar to

Query Match 63.3%; Score 26.6; DB 9; Length 172445;
Best Local Similarity 78.0%; Pred. No. 3.9;
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 GGGCCAGCCCTCTCTGAGAAATGACCATGGTTGACACAG 42
|||||
DB 42977 GGCCAGCTCTCTCTGAGCAATGCCCATGATACACAG 43017
|||||

RESULT 6
HSPFARGAM HSPFARGAM 1766 bp mRNA linear PRI 21-AUG-1997
LOCUS H.sapiens mRNA for peroxisome proliferator activated receptor
DEFINITION gamma.
ACCESSION X90563
VERSION X90563.1 GI:1480099

KEYWORDS peroxisome proliferator-activated receptor gamma.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1766)
AUTHORS Lambe, K.G. and Tugwood, J.D.
TITLE A human peroxisome-proliferator-activated receptor-gamma is
activated by inducers of adipogenesis, including thiazolidinedione
drugs
JOURNAL Eur. J. Biochem. 239 (1), 1-7 (1996)
MEDLINE 96305359
PUBMED 8706692
REFERENCE 2 (bases 1 to 1766)
AUTHORS Lambe, K.G.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1995) K.G. Lambe, Zeneca Central Toxicology Lab.,
Alderley Park, Macclesfield, Cheshire SK10 4TG, UK

FEATURES

Location/Qualifiers
1..1766
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CL15-2a-2"
/tissue_type="placenta"
/clone_lib="lambda ZAP II"
103..1536
/note="alternative start site"
/codon_start=1
/product="peroxisome proliferator activated receptor
gamma"
/protein_id="CAA62153.1"
/db_xref="GI:1490313"
/db_xref="SWISS-PROT:P37231"
/translation="MTVDTMPFNTFGISSVDLSVMDHSHSDIKPFTTVDSSIS
ISTPHVEDIPRTDPVADYKIDLKQYSAIKVEPASPYPYSEKTLKPKPHEEP
SNSLMAIECRVCGDKASGHYGVHACGCKGFFRTIRLKYDRCDLNCRIHKHSPN
KQCYCRFOKLVAGSHNAIRFGMPQAEKELAEISSDIQNPESADLALAKHL
YDSYSPPLTKAKARAILTKTDSKSPVYDMSLMGMDKIKPKHITPQEOSKEVA
VAIRFQCGQPSRSEAVQETTEYAKSIPGVNLDNDQVTLKYGVEHIIYTMASLM
NKGVLISEGGQWTRFELKLRKPGDPMKPEFAVKNALELDDSLAIFIAVII
LSDRPGLLNVKPIEDIQNLQALELQKLNHPESSQLFAKLLQKMTDLRQIVTEHV
QLLQVKKTTDMSLHPLLEIYKDY"
109..1536
/note="alternative start site"
/codon_start=1
/product="peroxisome proliferator activated receptor
gamma"
/protein_id="CAA62153.1"
/db_xref="GI:1490314"
/db_xref="SWISS-PROT:P37231"
/translation="MTVDTMPFNTFGISSVDLSVMDHSHSDIKPFTTVDSSIS
TPHVEDIPRTDPVADYKIDLKQYSAIKVEPASPYPYSEKTLKPKPHEEP
SNSLMAIECRVCGDKASGHYGVHACGCKGFFRTIRLKYDRCDLNCRIHKHSPN
KQCYCRFOKLVAGSHNAIRFGMPQAEKELAEISSDIQNPESADLALAKHL
YDSYSPPLTKAKARAILTKTDSKSPVYDMSLMGMDKIKPKHITPQEOSKEVA
VAIRFQCGQPSRSEAVQETTEYAKSIPGVNLDNDQVTLKYGVEHIIYTMASLM
NKGVLISEGGQWTRFELKLRKPGDPMKPEFAVKNALELDDSLAIFIAVII
LSDRPGLLNVKPIEDIQNLQALELQKLNHPESSQLFAKLLQKMTDLRQIVTEHV
QLLQVKKTTDMSLHPLLEIYKDY"

CDS

BASE COUNT 530 a 414 c 397 g 425 t
ORIGIN
Query Match 61.9%; Score 26; DB 9; Length 1766;
Best Local Similarity 85.3%; Pred. No. 10;
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 GCGCTCTCTCAGAAATGACCATGGTTGACACAG 42
|||||
DB 88 GCGGTGCCCCAGAAATGACCATGGTTGACACAG 121
|||||

RESULT 7
HUMPFARGE

LOCUS HUMPPARG 1808 bp mRNA linear PRI 26-DEC-2001
 DEFINITION Homo sapiens peroxisome proliferator activated receptor gamma (PPARG) mRNA, complete cds.
 ACCESSION L40904
 VERSION L40904.2 GI:17978515
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1808)
 AUTHORS Greene, M.E., Blumberg, B., McBride, O.W., Yi, H.F., Kronquist, K., Kwan, K., Hsieh, L., Greene, G. and Nimer, S.D.
 TITLE Isolation of the human peroxisome proliferator activated receptor gamma cDNA: expression in hematopoietic cells and chromosomal mapping
 JOURNAL Gene Expr. 4 (4-5), 281-299 (1995)
 MEDLINE 95107078
 PUBMED 7787419
 REFERENCE 2 (bases 1 to 1808)
 AUTHORS Qi, J.S., Desai-Yajnik, V., Greene, M.E., Raaka, B.M. and Samuels, H.H.
 TITLE The ligand-binding domains of the thyroid hormone/receptor gene subfamily function in vivo to mediate heterodimerization, gene silencing, and transactivation
 JOURNAL Mol. Cell. Biol. 15 (3), 1817-1825 (1995)
 MEDLINE 9316267
 PUBMED 7862171
 REFERENCE 3 (bases 1 to 1808)
 AUTHORS Greene, M.E., Blumberg, B., McBride, O.W., Yi, H.F., Kronquist, K., Kwan, K., Hsieh, L., Greene, G. and Nimer, S.D.
 TITLE Direct Submission
 JOURNAL Submitted (09-JAN-1991) Department of Medicine, UCLA School of Medicine 90024, USA
 REFERENCE 4 (bases 1 to 1808)
 AUTHORS Greene, M.E.
 TITLE Direct Submission
 JOURNAL Submitted (26-DEC-2001) University of Chicago, Chicago, IL 60637, USA
 REMARK Sequence update by submitter
 COMMENT On Dec 26, 2001 this sequence: version replaced gi:722619.
 FEATURES
 source
 1..1808
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3"
 /map="3p25"
 /clone="14"
 /sex="female"
 /tissue_type="bone marrow aspirate"
 /dev_stage="adult"
 1..1808
 /gene="PPARG"
 /note="h ppar gamma gene"
 173..1606
 /gene="PPARG"
 /note="ligand-activated transcription factor; nuclear receptor"
 /codon_start=1
 /product="peroxisome proliferator activated receptor gamma"
 /protein_id="AAA80314.2"
 /db_xref="GI:17978516"
 /translation="MTWVDTEMPFTNFGISSVDLSWEDHSHSDIKPFTTVDFSS
 ISTPHYEDIPRTDPVADYKYLQYGAIKVEPASPPIYSEKTLQYNKPHPEP
 SNLMAICRGKAGSGFHYGVHAGECKGFRTIRLKIYDRCDLNCRIHKSRN
 KCQCRFKCLAVGMSNARFGFMFOAEKELAEISSIDOLNPESADLALAKHL
 VSYIKSLPKAKARLLTKTWDKSPFIYDMNSLMWGEDTKPKKHITPLQEOSKE
 NKGVLISGEGQFMTRFSLKSRPFGDFMEKFEFAVKFNALDSDLDLAIPIAVII
 LSGDRPGLNLVKKPFDIQNLQLQALQKLNHPSSQLFAKLQKMTDLRQIVTEHY
 QLLQVIKKTETDMSLHPLQLEIYKLY"
 1808
 /gene="PPARG"

BASE COUNT 509 a 432 c 422 g 445 t
 ORIGIN
 Query Match 61.9%; Score 26; DB 9; Length 1808;
 Best Local Similarity 85.3%; Pred. No. 10;
 Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 9 GCCCTCTCTCAGAAATGACCATGTTGACACAG 42
 |||||
 Db 158 GCCGTGCCGCGAGAAATGACCATGTTGACACAG 191
 RESULT 8
 AX409565 1811 bp DNA linear PAT 14-JUN-2002
 LOCUS
 DEFINITION Sequence 2212 from Patent WO0229103.
 ACCESSION AX409565
 VERSION AX409565.1 GI:21442270
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
 TITLE Gene expression profiles in liver cancer
 JOURNAL Patent: WO 0229103-A 2212 11-APR-2002;
 GENE LOGIC INC (US)
 FEATURES
 Location/Qualifiers
 source
 1..1811
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="EMBL/GenBank Accession No. L40904"
 BASE COUNT 510 a 433 c 422 g 446 t
 ORIGIN
 Query Match 61.9%; Score 26; DB 6; Length 1811;
 Best Local Similarity 85.3%; Pred. No. 10;
 Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 9 GCCCTCTCTCAGAAATGACCATGTTGACACAG 42
 |||||
 Db 158 GCCGTGCCGCGAGAAATGACCATGTTGACACAG 191
 RESULT 9
 ARI39020 1844 bp DNA linear PAT 16-JUN-2001
 LOCUS
 DEFINITION Sequence 1 from patent US 6200802.
 ACCESSION ARI39020
 VERSION ARI39020.1 GI:14481365
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1844)
 AUTHORS Greene, M.E. and Blumberg, B.
 TITLE Human peroxisome proliferator activated receptor gamma: compositions and methods
 JOURNAL Patent: US 6200802-A 1 13-MAR-2001;
 FEATURES
 Location/Qualifiers
 source
 1..1844
 /organism="unknown"
 BASE COUNT 543 a 433 c 421 g 447 t
 ORIGIN
 Query Match 61.9%; Score 26; DB 6; Length 1844;
 Best Local Similarity 85.3%; Pred. No. 10;
 Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 9 GCCCTCTCTCAGAAATGACCATGTTGACACAG 42
 |||||
 Db 158 GCCGTGCCGCGAGAAATGACCATGTTGACACAG 191

polyA_site


```

RESULT 10
AY048697
LOCUS
DEFINITION
Macaca fascicularis peroxisome proliferator-activated receptor
gamma 4 (PPARGgamma4) mRNA, complete cds.
ACCESSION
AY048697
VERSION
AY048697.1 GI:21552434
KEYWORDS
crab-eating macaque.
SOURCE
Macaca fascicularis
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopitheinae; Macaca.
REFERENCE
1 (bases 1 to 2063)
Zhou, J., Wilson, K.M. and Medh, J.D.
Genetic analysis of four novel peroxisome proliferator activated
receptor-gamma splice variants in monkey macrophages
Biochem. Biophys. Res. Commun. 293 (1), 274-283 (2002)
JOURNAL
22050014
MEDLINE
12054596
PUBMED
2 (bases 1 to 2063)
Zhou, J., Wilson, K.M. and Medh, J.D.
Direct Submission
Submitted (27-JUL-2001) Internal Medicine, The University of Iowa,
200 Hawkins Drive, Iowa, IA 52242, USA
JOURNAL
FEATURES
Location/Qualifiers
source
1..2063
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/cell_type="macrophage"
1..2063
/gene="PPARGgamma4"
354..1805
/gene="PPARGgamma4"
354..1805
/gene="PPARGgamma4"
/feature="transcription factor"
/feature="transcription factor"
/codon_start=1
/product="peroxisome proliferator-activated receptor gamma
4"
/protein_id="AAL05267.1"
/db_xref="GI:21552435"
/translation="MKGFSEITMVDTPMPNFGISSVDSVMDHSHSFDIKPPT
TVDFSSISAPHYEDIPFTPTDPMVADYKYLQEQSAIKVEPASPYPYSEKTLQYN
KPHEEPSNLMAIECVCGDKASGFHYGVHACGCKGFRRTIRKLIVDRCDLNCRI
HKSRNKCQVCFQKCLAVGSHNAIRFGMPQAEKEKLARISSDIDOLNPSADLR
ALAKHLYDSYKSPFLTKAKARAILTGKTDKSPFVIVMNSLWMDGDKYKPHIITPL
QEOKSEVAIRIFQGCQPSVEAOEITEYAKSIPGVNLDNDQVTLKYGVHEIYT
MLASLNKQGVFISSEGQFMTREFLKSRPFGDFMPEKPEFAVKFNALDSDLA
FIATVILSGDRPGLNVKPIEDIQNLQALEQLKLNHPSSQLFAXLLQKMTDLRQ
IVTEHVQLLVQIKTKTDMSLHPLLEIYKDLV"
BASE COUNT 623 a 452 c 438 g 515 t
ORIGIN
Query Match 57.6%; Score 24.2; DB 9; Length 2063;
Best Local Similarity 89.7%; Pred. No. 53;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 14 CCTCTCAGAAATGACCATGGTTGACACAG 42
Db 362 CTTTCAGAAATACCATGGTTGACACAG 390
RESULT 11
AY048699
LOCUS
DEFINITION
Macaca fascicularis peroxisome proliferator-activated receptor
gamma 6 (PPARGgamma6) mRNA, complete cds.
ACCESSION
AY048699
VERSION
AY048699.1 GI:21552438
KEYWORDS
crab-eating macaque.
SOURCE
Macaca fascicularis
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopitheinae; Macaca.
REFERENCE
1 (bases 1 to 2063)
Zhou, J., Wilson, K.M. and Medh, J.D.
Genetic analysis of four novel peroxisome proliferator activated
receptor-gamma splice variants in monkey macrophages
Biochem. Biophys. Res. Commun. 293 (1), 274-283 (2002)
JOURNAL
22050014
MEDLINE
12054596
PUBMED
2 (bases 1 to 2063)
Zhou, J., Wilson, K.M. and Medh, J.D.
Direct Submission
Submitted (27-JUL-2001) Internal Medicine, The University of Iowa,
200 Hawkins Drive, Iowa, IA 52242, USA
JOURNAL
FEATURES
Location/Qualifiers
source
1..2063
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/cell_type="macrophage"
1..2063
/gene="PPARGgamma6"
389..1840
/gene="PPARGgamma6"
/feature="transcription factor"
/feature="transcription factor"
/codon_start=1
/product="peroxisome proliferator-activated receptor gamma
6"
/protein_id="AAL05269.1"
/db_xref="GI:21552439"
/translation="MKGFSEITMVDTPMPNFGISSVDSVMDHSHSFDIKPPT
TVDFSSISAPHYEDIPFTPTDPMVADYKYLQEQSAIKVEPASPYPYSEKTLQYN
KPHEEPSNLMAIECVCGDKASGFHYGVHACGCKGFRRTIRKLIVDRCDLNCRI
HKSRNKCQVCFQKCLAVGSHNAIRFGMPQAEKEKLARISSDIDOLNPSADLR
ALAKHLYDSYKSPFLTKAKARAILTGKTDKSPFVIVMNSLWMDGDKYKPHIITPL
QEOKSEVAIRIFQGCQPSVEAOEITEYAKSIPGVNLDNDQVTLKYGVHEIYT
MLASLNKQGVFISSEGQFMTREFLKSRPFGDFMPEKPEFAVKFNALDSDLA
FIATVILSGDRPGLNVKPIEDIQNLQALEQLKLNHPSSQLFAXLLQKMTDLRQ
IVTEHVQLLVQIKTKTDMSLHPLLEIYKDLV"
BASE COUNT 637 a 460 c 440 g 526 t
ORIGIN
Query Match 57.6%; Score 24.2; DB 9; Length 2063;
Best Local Similarity 89.7%; Pred. No. 53;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 14 CCTCTCAGAAATGACCATGGTTGACACAG 42
Db 397 CTTTCAGAAATACCATGGTTGACACAG 425
RESULT 12
AY048699
LOCUS
DEFINITION
Macaca fascicularis peroxisome proliferator-activated receptor
gamma 6 (PPARGgamma6) mRNA, complete cds.
ACCESSION
AY048699
VERSION
AY048699.1 GI:21552438
KEYWORDS
crab-eating macaque.
SOURCE
Macaca fascicularis
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopitheinae; Macaca.
REFERENCE
1 (bases 1 to 135011)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
Alsbrooks, S.L., Amarutunge, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbaria, J., Benton, J., Bimaga, K., Blankenburg, K., Bonni, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,

```

Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
 Homsif, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kuresni, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtenberg, O., Lieu, C., Liu, J., Liu, W., Louisseghed, H.,
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, N., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Prinus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
 Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
 Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Teifrod, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczek, R., Wooden, S., Wolley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstock, G., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 135011)
 Worley, K.C.
 Direct Submission
 Submitted (24-NOV-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 135011)
 Worley, K.C.
 Direct Submission
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 11, 2002 this sequence version replaced gi:17974449.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: G1ZS
 Center clone name: CH230-222E15
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 82376 bases at least Q40
 Consensus quality: 86006 bases at least Q30
 Consensus quality: 88830 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length.
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 52 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

* 68931 68930: gap of unknown length
* 71789: contig of 2859 bp in length
* 71889: gap of unknown length
* 71790 71889: contig of 5638 bp in length
* 77587 77587: gap of unknown length
* 77688 81108: contig of 3421 bp in length
* 81109 81208: gap of unknown length
* 84209 84806: contig of 3598 bp in length
* 84907 89112: contig of 4206 bp in length
* 89113 91218: gap of unknown length
* 89213 93228: gap of unknown length
* 93129 96338: contig of 3110 bp in length
* 96339 96438: gap of unknown length
* 96439 100306: contig of 3868 bp in length
* 100307 100406: gap of unknown length
* 100407 104867: contig of 4461 bp in length
* 104868 104967: gap of unknown length
* 108910 109010: contig of 3943 bp in length
* 108911 114113: contig of 5103 bp in length
* 114114 114213: gap of unknown length
* 114214 117375: contig of 3162 bp in length
* 117376 117475: gap of unknown length
* 117476 120862: contig of 3387 bp in length
* 120863 120963: gap of unknown length
* 120963 127749: contig of 6787 bp in length
* 127750 127850 135011: contig of 7162 bp in length.

```

FEATURES

Location/Qualifiers
 1..135011
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"

Query Match 56.78; Score 23.8; DB 2; Length 135011;
 Best Local Similarity 92.63; Pred. No. 53;
 Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 CCAGGCCCTCTCTCAGAAATGACCAT 31
 |||||
 DB 121477 CCAGGCCCTCTCTCAGAAATGACCAT 121503

RESULT 13
 AC112596
 LOCUS Rattus norvegicus clone CH230-45P12, *** SEQUENCING IN PROGRESS
 DEFINITION *** 55 unordered pieces.
 AC112596
 VERSION AC112596.3 GI:21743150
 KEYWORDS HTG; HTGS_PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 182241)

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, P.R., Allen, C.,
 Alsbrooks, S.L., Amaral, H.C., Are, J.R., Ayele, M., Banks, P.,
 Barbara, J., Benton, J., Blum, K., Blum, K., Blum, K., Blum, K.,
 Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
 Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Deigado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,

Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
 Honsi, P., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokwenkwo, S., Ogih, M., Okwuonu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojupokan, I., Rolfe, M., Ruiz, S., Savary, G.,
 Scherer, S., Scott, G., Sher, H., Shooshtari, N., Sisson, I.,
 Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstock, G. and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 182241)
 Worley, K.C.
 Direct Submission
 Submitted (22-FEB-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 182241)
 Worley, K.C.
 Direct Submission
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 12, 2002 this sequence version replaced gi:20303210.

COMMENT

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GFEN
 Center clone name: CH230-45P12
 ----- Summary Statistics
 Sequencing vector: Plasmid.
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990129
 Consensus quality: 132458 bases at least Q40
 Consensus quality: 136923 bases at least Q30
 Consensus quality: 140922 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 55 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1421: contig of 1421 bp in length
 * 1422 1521: gap of unknown length
 * 1522 2743: contig of 1222 bp in length
 * 2744 2843: gap of unknown length
 * 2844 3897: contig of 1054 bp in length
 * 3898 3997: gap of unknown length
 * 3998 5369: contig of 1372 bp in length

* 5370 5469: gap of unknown length
 * 5470 7029: contig of 1560 bp in length
 * 7030 7129: gap of unknown length
 * 7130 9029: contig of 1900 bp in length
 * 9030 9129: gap of unknown length
 * 9130 10151: contig of 1022 bp in length
 * 10152 10251: gap of unknown length
 * 10252 11287: contig of 1036 bp in length
 * 11288 11387: gap of unknown length
 * 11388 13276: contig of 1889 bp in length
 * 13277 13376: gap of unknown length
 * 13377 14699: contig of 1323 bp in length
 * 14700 14799: gap of unknown length
 * 14800 16353: contig of 1554 bp in length
 * 16354 16453: gap of unknown length
 * 16454 18403: contig of 1950 bp in length
 * 18404 18503: gap of unknown length
 * 18504 19527: contig of 1024 bp in length
 * 19528 19627: gap of unknown length
 * 19628 20969: contig of 1342 bp in length
 * 20970 21069: gap of unknown length
 * 21070 22706: contig of 1637 bp in length
 * 22707 22806: gap of unknown length
 * 22807 24106: contig of 1300 bp in length
 * 24107 24206: gap of unknown length
 * 24207 25722: contig of 1516 bp in length
 * 25723 25822: gap of unknown length
 * 25823 27934: contig of 2112 bp in length
 * 27935 28034: gap of unknown length
 * 28035 30821: contig of 2787 bp in length
 * 30822 30921: gap of unknown length
 * 30922 32675: contig of 1754 bp in length
 * 32676 32775: gap of unknown length
 * 32776 35269: contig of 2494 bp in length
 * 35270 35369: gap of unknown length
 * 35370 37118: contig of 1749 bp in length
 * 37119 37218: gap of unknown length
 * 37219 39752: contig of 2534 bp in length
 * 39753 39852: gap of unknown length
 * 39853 40888: contig of 1036 bp in length
 * 40889 40988: gap of unknown length
 * 40989 44479: contig of 3491 bp in length
 * 44480 44579: gap of unknown length
 * 44580 47930: contig of 3351 bp in length
 * 47931 48030: gap of unknown length
 * 48031 51350: contig of 3320 bp in length
 * 51351 51450: gap of unknown length
 * 51451 53819: contig of 2369 bp in length
 * 53820 53919: gap of unknown length
 * 53920 55492: contig of 1573 bp in length
 * 55493 55592: gap of unknown length
 * 55593 57253: contig of 1661 bp in length
 * 57254 57353: gap of unknown length
 * 57354 59139: contig of 1786 bp in length
 * 59140 59239: gap of unknown length
 * 59240 61621: contig of 2382 bp in length
 * 61622 61721: gap of unknown length
 * 61722 63910: contig of 2189 bp in length
 * 63911 64010: gap of unknown length
 * 64011 66403: contig of 2393 bp in length
 * 66404 66503: gap of unknown length
 * 66504 69190: contig of 2687 bp in length
 * 69191 69290: gap of unknown length
 * 69291 73197: contig of 3907 bp in length
 * 73198 73297: gap of unknown length
 * 73298 77538: contig of 4241 bp in length
 * 77539 80543: contig of 2905 bp in length
 * 80544 80644: gap of unknown length
 * 80644 83399: contig of 2756 bp in length
 * 83400 83499: gap of unknown length
 * 83500 87433: contig of 3934 bp in length
 * 87434 87533: gap of unknown length

* 87534 92524: contig of 4991 bp in length
 * 92525 92624: gap of unknown length
 * 92625 96337: contig of 3713 bp in length
 * 96338 96437: gap of unknown length
 * 96438 100807: contig of 4370 bp in length
 * 100808 100907: gap of unknown length
 * 100908 105330: contig of 4423 bp in length
 * 105331 105430: gap of unknown length
 * 105431 109318: contig of 3888 bp in length
 * 109319 109418: gap of unknown length
 * 109419 113116: contig of 3698 bp in length
 * 113117 113216: gap of unknown length
 * 113217 117509: contig of 4293 bp in length
 * 117510 117609: gap of unknown length
 * 117610 124351: contig of 6741 bp in length
 * 124351 124450: gap of unknown length
 * 124451 129595: contig of 5145 bp in length
 * 129596 129695: gap of unknown length
 * 129696 136094: contig of 6399 bp in length
 * 136095 136194: gap of unknown length
 * 136195 143961: contig of 7766 bp in length
 * 143961 144060: gap of unknown length
 * 144061 150311: contig of 6251 bp in length
 * 150312 150411: gap of unknown length
 * 150412 157843: contig of 7432 bp in length
 * 157844 157943: gap of unknown length
 * 157944 167965: contig of 10022 bp in length

Query Match 56.7%; Score 23.8; DB 2; Length 182241;
 Best Local Similarity 92.6%; Pred. No. 52;
 Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 CCAGGCGCTCTCTCAGAAATGACCAT 31
 |||||
 Db 3033 CCAGGCGCTCTCTCAGAAATGACCAT 3059

RESULT 14
 AC004387/c
 LOCUS AC004387 206156 bp DNA linear HTG 13-JUN-2002
 DEFINITION Homo sapiens chromosome X clone BWXD173, *** SEQUENCING IN PROGRESS
 *** 2 ordered pieces.
 ACCESSION AC004387
 VERSION AC004387.2 GI:21405643
 KEYWORDS HTG; HTGS PHASE2.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 206156)
 AUTHORS Chen, R., Brownstein, B.H., States, D.J., Schlessinger, D. and
 Mazzarella, R.
 TITLE Direct Submission
 JOURNAL Unpublished (1997)
 REFERENCE 2 (bases 1 to 206156)
 AUTHORS Brownstein, B.H., States, D.J. and Mazzarella, R.
 TITLE Direct Submission
 JOURNAL Submitted (09-MAR-1998) Center for Genetics in Medicine, Box 8232,
 Washington University School of Medicine, 4566 Scott Avenue, St.
 Louis, MO 63110, USA
 COMMENT On Jun 13, 2002 this sequence version replaced gi:2944106.
 Current status of this project is available at:
 'http://www.ibc.wustl.edu/cgm/seq_projects.html'.
 Submitted by:
 Ellison Chen,
 Advanced Center for Genetic Technology,
 Applied Biosystems Division of Perlin Elmer Corp.,
 850 Lincoln Center Drive,
 Foster City, CA 94404 USA
 e-mail: ellison@genseq.apltdbio.com
 and

Buddy Brownstein,
Center for Genetics in Medicine,
Washington University School of Medicine, Box 8232
4566 Scott Avenue,
St. Louis, MO 63110, USA
e-mail: buddy@genetics.wustl.edu

and

David J. States,
Institute for Biomedical Computing
Washington University in St. Louis
700 South Euclid Ave.
St. Louis, MO 63108 USA
e-mail: states@bc.wustl.edu

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* I 195298: contig of 195298 bp in length
* 195299 195398: gap of 100 bp
* 195399 206156: contig of 10758 bp in length.

FEATURES
Source
1..206156
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="bWXD173"

BASE COUNT 58189 a 43613 c 43696 g 60558 t 100 others
ORIGIN

Query Match 56.2%; Score 23.6; DB 2; Length 206156;
Best Local Similarity 76.3%; Pred. No. 62;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 GGCCAGGCCCTCTCTCAGAAATGACCATGGTTGACA 39

Db 138897 GTCCAGGCACTCTTCTTAAGCAGTGACCAAGGCGACGA 138860

RESULT 15
AL672032/c

LOCUS Human DNA sequence from clone RP11-308B5 on chromosome X, complete
DI TION AL672032 209317 bp DNA linear PRI 02-MAY-2002

ACCESSION AL672032
VERSION AL672032.6 GI:20428914

KEYWORDS HTG

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Whitehead, S.

Direct Submission

Submitted (01-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk

On May 3, 2002 this sequence version replaced gi:20386906.
During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone. This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/Celegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX

RP11-308B5 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6.

FEATURES
Source

1..209317
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-308B5"
/clone_lib="RPCI-11.2"

BASE COUNT 62375 a 44037 c 43724 g 59181 t

ORIGIN

Query Match 56.2%; Score 23.6; DB 9; Length 209317;
Best Local Similarity 76.3%; Pred. No. 62;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 GGCCAGGCCCTCTCTCAGAAATGACCATGGTTGACA 39

Db 39518 GTCCAGGCACTCTTCTTAAGCAGTGACCAAGGCGACGA 39481

Search completed: January 21, 2003, 23:57:13
Job time : 1017.33 secs

DR WPI: 2001-514487/56.
 XX P-PSDB: AAB85797.
 PT New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
 PT treating certain tumors or cancers, e.g. follicular carcinoma -
 XX
 PS Claim 21; Page 118; 145pp; English.
 XX
 CC The invention relates to an oncogene designated PAX8-PPARGamma1 that
 CC contains a PAX8 coding region fused to PPARGamma1 coding region. The
 CC PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant
 CC methodology. A PPARGamma1 ligand or agent is useful for treating a
 CC subject having a disorder characterized by the presence of a PAX8-
 CC PPARGamma1, where the disorder is cancer, e.g. follicular carcinoma.
 CC The PAX8-PPARGamma1 molecules are also useful for providing nucleotide
 CC and amino acid sequences useful for detecting the above disease. The
 CC present sequence represents the nucleotide sequence including and
 CC surrounding the fusion junction in the PAX8e8-PPARGamma1 DNA.
 XX
 SQ Sequence 42 BP; 10 A; 14 C; 11 G; 7 T; 0 other;
 Query Match 100.0%; Score 42; DB 22; Length 42;
 Best Local Similarity 100.0%; Pred. No. 3.2e-08;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGGGCCAGGCCCTCTCTCAGAAATGACCATGGTTGACACAG 42
 DB 1 CGGGCCAGGCCCTCTCTCAGAAATGACCATGGTTGACACAG 42
 RESULT 2
 AAH76282
 ID AAH76282 standard; CDNA; 2523 BP.
 XX
 AC AAH76282;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Human PAX8e8-PPARGamma1 cDNA sequence.
 XX
 KW PAX8-PPARGamma1; oncogene; cytostatic; PAX8; PPARGamma1; cancer;
 KW follicular carcinoma; PAX8e8-PPARGamma1; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2523
 FT /*tag= a
 XX
 PN WO200152789-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 18-JAN-2001; 2001WO-US01664.
 XX
 PR 20 JAN-2000; 2000US-0177109.
 DR 14-AUG-2000; 2000US-0225079.
 XX
 PA (BGM) BRIGHAM & WOMENS HOSPITAL INC.
 XX
 PI Kroll TG, Fletcher JA;
 XX
 DR WPI: 2001-514487/56.
 DR P-PSDB: AAB85794.
 XX
 CC New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
 PT treating certain tumors or cancers, e.g. follicular carcinoma -
 XX
 PS Claim 1; Page 106-109; 145pp; English.
 XX
 CC The invention relates to an oncogene designated PAX8-PPARGamma1 that
 CC contains a PAX8 coding region fused to PPARGamma1 coding region. The
 CC PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant

CC methodology. A PPARGamma1 ligand or agent is useful for treating a
 CC subject having a disorder characterized by the presence of a PAX8-
 CC PPARGamma1, where the disorder is cancer, e.g. follicular carcinoma.
 CC The PAX8-PPARGamma1 molecules are also useful for providing nucleotide
 CC and amino acid sequences useful for detecting the above disease. The
 CC present sequence represents a human PAX8e8-PPARGamma1 polypeptide
 XX
 SQ Sequence 2523 BP; 643 A; 723 C; 614 G; 543 T; 0 other;
 Query Match 100.0%; Score 42; DB 22; Length 2523;
 Best Local Similarity 100.0%; Pred. No. 8.2e-08;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGGGCCAGGCCCTCTCTCAGAAATGACCATGGTTGACACAG 42
 DB 1067 CGGGCCAGGCCCTCTCTCAGAAATGACCATGGTTGACACAG 1108
 RESULT 3
 AAX19066
 ID AAX19066 standard; DNA; 695 BP.
 XX
 AC AAX19066;
 XX
 DT 13-MAY-1999 (first entry)
 XX
 DE Human PPARG-gamma intron B, exon 1 and intron 1.
 XX
 KW Human; peroxisome proliferator activated receptor gamma; PPARG-gamma;
 KW regulatory sequence; promoter; obesity; anorexia; lipoma; cachexia;
 KW lipodystrophy; liposarcoma; human immunodeficiency virus; HIV;
 KW insulin resistance; non-insulin-dependent diabetes mellitus;
 KW polycystic ovary syndrome; gastrointestinal tract; Crohn's disease;
 KW inflammatory bowel disease; ulcerative colitis; bowel cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9905161-A1.
 XX
 PD 04-FEB-1999.
 XX
 PF 24-JUL-1998; 98WO-US15411.
 XX
 PR 25-JUL-1997; 97US-0053692.
 XX
 PA (INSP) INST PASTEUR.
 PA (LIGA-) LIGAND PHARM INC.
 XX
 PI Auwerx J, Briggs MR, Fajas L, Saladin RS;
 XX
 DR WPI: 1999-142844/12.
 XX
 CC Newly isolated nucleic acid comprising a control region of a human
 PT peroxisome proliferator activated receptor (PPAR) gamma gene -
 PT useful for identifying modulators that are useful in treating
 PT diseases associated with abnormal levels of human PPARG-gamma gene
 PT expression
 XX
 PS Disclosure; Page 89; 102pp; English.
 XX
 CC The present invention describes an isolated, purified or enriched
 CC nucleic acid comprising a control region of a human peroxisome
 CC proliferator activated receptor gamma (PPAR-gamma) gene. The nucleic
 CC acids are useful for screening for agents capable of modulating the
 CC expression of a human PPARG-gamma gene. These agents (modulators) form
 CC pharmaceutical compositions that are useful for treating diseases
 CC associated with high/low levels of human PPARG-gamma gene expression.
 CC The diseases include obesity, anorexia, cachexia, lipodystrophy, lipomas,
 CC liposarcomas, abnormalities associated with anti-human immunodeficiency
 CC virus (HIV) treatment, insulin resistance, non-insulin-dependent diabetes
 CC mellitus (NIDDM), polycystic ovary syndrome, diseases of the
 CC gastrointestinal (GI) tract, inflammatory bowel disease, Crohn's disease,

CC ulcerative colitis and bowel cancer. The nucleic acids are useful for
 CC studying the role of the PPAR-gamma gene in various diseases and
 CC disorders. The structure of PPAR-gamma enables genetic studies of PPAR-
 CC gamma mutations in humans, and evaluation of its role in disorders like
 CC insulin resistance, NIDDM, and diseases associated with altered adipose
 CC tissue function, like obesity and lipodystrophic syndromes. The nucleic
 CC acids are also useful for gene therapy and the production of transgenic
 CC animals, which are useful in screening assays. The control regions of
 CC the nucleic acids enable screening for modulators of the human PPAR-gamma
 CC gene, which are useful in designing drugs for treating disorders or
 CC diseases associated with the level of PPAR-gamma gene expression.
 CC The present sequence represents the human PPAR-gamma intron B, exon 1
 CC and intron 1.

XX SQ Sequence 695 BP; 196 A; 148 C; 130 G; 221 T; 0 other;

Query Match 67.6%; Score 28.4; DB 20; Length 595;
 Best Local Similarity 96.7%; Pred. No. 0.035; Indels 0; Gaps 0;
 Matches 29; Conservative 0; Mismatches 1;

Qy 13 TCCTCTCAGAAATGACCATGTTGACACAG 42
 |||||
 Db 424 TCCTTCAGAAATGACCATGTTGACACAG 453

RESULT 4

AA36522
 ID AAX36522 standard; cDNA; 1811 BP.

AC AAX36522;

DT 07-JUL-1999 (first entry)

XX Human PPAR-gamma coding sequence.

XX Nuclear receptor agonist; antagonist; identification; PPAR;
 XX peroxisome proliferator activated receptor; ss.

XX Homo sapiens.

XX WO9918124-A1.

XX 15-APR-1999.

XX 06-OCT-1998; 98WO-US21049.

XX 07-OCT-1997; 97US-0061385.

XX 'MERI' MERCK & CO INC.

XX Cummings RT, Hermes JD, Moller DE, Zhou G;

XX WPI; 1999-263998/22.

XX P-PSDB; AAY05471.

XX Identifying nuclear receptor agonists and antagonists

XX Disclosure; Fig 9b; 60pp; English.

XX This sequence encodes the human peroxisome proliferator activated
 XX receptor-gamma (PPAR-gamma).

XX The invention relates to a method for identifying nuclear receptor
 XX agonists and antagonists comprises measuring fluorescent resonance energy
 XX transfer between fluorescent-labelled nuclear receptors and
 XX co-activators. The method can be used for identifying agonists and
 XX antagonist of nuclear receptors.

XX SQ Sequence 1811 BP; 510 A; 433 C; 422 G; 446 T; 0 other;

Query Match

Best Local Similarity 61.9%; Score 26; DB 20; Length 1811;
 Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 9 GCCCTCCTCTCAGAAATGACCATGTTGACACAG 42
 |||||
 Db 158 GCCGTGCGCGCAGAAATGACCATGTTGACACAG 191

RESULT 5

AAH76296

ID AAH76296 standard; cDNA; 1811 BP.

AC AAH76296;

XX 29-OCT-2001 (first entry)

XX Human PPARgamma cDNA sequence.

XX PAX8-PPARGamma; oncogene; cytostatic; PAX8; PPARGamma; cancer;
 XX follicular carcinoma; PPARGamma; human; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 173...1609

FT /*tag= a

XX PN WO200152789-A2.

XX 26-JUL-2001.

XX 18-JAN-2001; 2001WO-US01664.

XX 20-JAN-2000; 2000US-0177109.

XX 14-AUG-2000; 2000US-0225079.

XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.

XX Kroll TG, Fletcher JA;

XX WPI; 2001-514487/56.

XX P-PSDB; AAB85802.

XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
 XX treating certain tumors or cancers, e.g. follicular carcinoma

XX Disclosure; Page 137-139; 145pp; English.

XX The invention relates to an oncogene designated PAX8-PPARGamma that
 CC contains a PAX8 coding region fused to PPARGamma coding region. The
 CC PAX8-PPARGamma polypeptides can be expressed by standard recombinant
 CC methodology. A PPARGamma ligand or agent is useful for treating a
 CC subject having a disorder characterized by the presence of a PAX8-
 CC PPARGamma, where the disorder is cancer, e.g. follicular carcinoma.
 CC The PAX8-PPARGamma molecules are also useful for providing nucleotide
 CC and amino acid sequences useful for detecting the above disease. The
 CC present sequence represents a human PPARGamma polypeptide encoding cDNA.

XX SQ Sequence 1811 BP; 510 A; 433 C; 422 G; 446 T; 0 other;

Query Match

Best Local Similarity 61.9%; Score 26; DB 22; Length 1811;
 Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 9 GCCCTCCTCTCAGAAATGACCATGTTGACACAG 42
 |||||
 Db 158 GCCGTGCGCGCAGAAATGACCATGTTGACACAG 191

RESULT 6

ABN95714

ID ABN95714 standard; DNA; 1811 BP.

XX ABN95714;

XX 13-AUG-2002 (first entry)

XX DE Gene #2212 used to diagnose liver cancer.
 XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 KW KW metastatic liver tumor; cytostatic; expression profile; disease state;
 KW KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX OS Homo sapiens.
 XX XX WO200229103-A2.
 XX PN 11-APR-2002.
 XX PD 02-OCT-2001; 2001WO-US30589.
 XX XX 02-OCT-2000; 2000US-237054P.
 XX PR (GENE-) GENE LOGIC INC.
 XX PA Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 XX PI WPI; 2002-426119/45.
 XX DR Diagnosing and detecting the progression of liver cancer.
 XX PT hepatocellular carcinoma or metastatic liver tumor in a patient, in a
 PT PT involves detecting the level of expression of two or more genes in a
 PT PT liver tissue sample.
 XX PS Claim 1; SEQ ID NO 2212; 298pp; English.
 XX CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumor in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 1811 BP; 510 A; 433 C; 422 G; 446 T; 0 other;
 Query Match 61.9%; Score 26; DB 24; Length 1811;
 Best Local Similarity 85.3%; Pred. No. 0.45;
 Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 9 GCCCTCCTCTCAGAAATGACCATGCTTGACACAG 42
 DB 158 GCCGTGGCGGAGAAATGACCATGCTTGACACAG 191
 RESULT 7
 AAF55663
 ID AAF55663 standard; DNA; 1844 BP.
 XX AC AAF55663;
 XX DT 29-MAY-2001 (first entry)
 XX DE Human peroxisome proliferator-activated receptor gamma DNA.
 XX KW Human; peroxisome proliferator-activated receptor gamma; PPAR-gamma;
 KW KW orphan receptor; cancer; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers

FT CDS 179..1609
 FT FT /*tag= a
 FT FT /product= "peroxisome proliferator-activated receptor
 XX gamma"
 XX PN US6200802-B1.
 XX PD 13-MAR-2001.
 XX XX 08-OCT-1993; 93US-0134557.
 XX PF 08-OCT-1993; 93US-0134557.
 XX PR (ARCH-) ARCH DEV CORP.
 XX PA Greene ME, Blumberg B;
 XX PI WPI; 2001-234517/24.
 XX D? P-PSDB; AAB67656.
 XX SQ Novel isolated and purified polynucleotide encoding human peroxisome
 PT PT proliferator-activated receptor gamma polypeptide useful in
 PT PT diagnostics, drug design and therapeutics
 XX PS Claim 3; Column 43-48; 29pp; English.
 XX CC The present sequence encodes a human peroxisome proliferator-activated
 CC receptor gamma (PPAR-gamma) polypeptide. PPAR are orphan receptors, and
 CC may play a role in proliferative and differentiation aspects of cancer.
 CC The PPAR-gamma polynucleotide is useful for detecting a mRNA transcript
 CC that encodes PPAR polypeptide. It is also useful for detecting hybrid
 CC formation. The PPAR-gamma polypeptide is useful in diagnostics, drug
 CC design and therapeutics.
 XX SQ Sequence 1844 BP; 543 A; 433 C; 421 G; 447 T; 0 other;
 Query Match 61.9%; Score 26; DB 22; Length 1844;
 Best Local Similarity 85.3%; Pred. No. 0.46;
 Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 9 GCCCTCCTCTCAGAAATGACCATGCTTGACACAG 42
 DB 158 GCCGTGGCGGAGAAATGACCATGCTTGACACAG 191
 RESULT 8
 AAS44653/c
 ID AAS44653 standard; DNA; 2295 BP.
 XX AC AAS44653;
 XX DT 18-DEC-2001 (first entry)
 XX DE Human full-length polynucleotide sequence #78.
 XX KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
 KW KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
 KW KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
 KW KW nervous system disorder; inflammatory disorder; cell differentiation; ds;
 KW KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
 KW KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
 KW KW cytostatic; antirheumatic; antiarthritic; vulnery; antinflammatory;
 KW KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
 KW KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
 KW KW immunostimulant; analgesic; gene therapy.
 XX OS Homo sapiens.
 XX PN WO200164834-A2.
 XX XX 07-SEP-2001.
 XX XX 26-FEB-2001; 2001WO-US04926.

```
XX 18-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
PR 17-JUN-2000; 2000US-0597107.
PR 14-JUL-2000; 2000US-0616807.
PR 19-SEP-2000; 2000US-0664641.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
PI Drmanac R;
XX WPI; 2001-589862/66.
DR P-PSDB; AAU27753.
XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis, treatment of
PT cancer, neurological, inflammatory disorders and for use in arrays for
PT detection -
XX Claim 1; SEQ ID No 78; 153pp; English.
XX Sequences AAS44576-AAS44919 represent full-length polynucleotides and
CC contig polynucleotides encoding polypeptides of the invention. The DNA
CC and protein sequences are useful for the treatment, diagnosis and
CC prevention of various types of disorder in a mammalian subject such as a
CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC bowel disease. The sequences exhibit activity relating to angiogenesis,
CC cell proliferation, cell differentiation, stem cell growth factor,
CC activin or inhibin. Therefore, they can be used to manipulate stem cells
CC in culture to give rise to neuroepithelial cells that can be used to
CC augment or replace cells damaged by illness, accidental damage or genetic
CC disorders. The sequences may also be used for regeneration of bone,
CC cartilage, tendons and ligaments and in tissue repair and burn healing.
CC Note: Some sequences for this patent did not form part of the printed
CC specification, but were obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 2295 BP; 525 A; 606 C; 565 G; 599 T; 0 other;
XX Y Match 61.9%; Score 26; DB 22; Length 2295;
XX Local Similarity 85.3%; Pred. No. 0.48;
XX Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
OY 9 GCCTCTCTCAGAAATGACCATGGTTGACACAG 42
DB 1654 GCCGTGGCGGAGAAATGACCATGGTTGACACAG 1621
RESULT 9
AAH76284
ID AAH76284 standard; DNA; 42 BP.
XX AAH76284;
XX 29-OCT-2001 (first entry)
XX Human PAX8e7-PPARGammael translocation fusion juncture 1 DNA sequence.
DE PAX8e7-PPARGammael translocation fusion juncture 1 DNA sequence.
DE PAX8-PPARGammael; oncogene; cytostatic; PAX8; PPARGammael; cancer;
KW follicular carcinoma; PAX8e7-PPARGammael; human; ss.
KW Homo sapiens.
OS Homo sapiens.
XX Key Location/Qualifiers
PH Key Location/Qualifiers
FT CDS 1..2334
FT /*tag= a
XX WO200152789-A2.
XX 26-JUL-2001.
XX 18-JAN-2001; 2001WO-US01664.
XX PAX8-PPARGammael; oncogene; cytostatic; PAX8; PPARGammael; cancer;
KW follicular carcinoma; PAX8e7-PPARGammael; human; ss.
KW Homo sapiens.
OS Homo sapiens.
XX Key Location/Qualifiers
PH Key Location/Qualifiers
FT CDS 1..2334
FT /*tag= a
XX WO200152789-A2.
XX 26-JUL-2001.
XX 18-JAN-2001; 2001WO-US01664.
XX 20-JAN-2000; 2000US-0177109.
```

```
FT CDS 3..41
FT /*tag= a
XX WO200152789-A2.
XX 26-JUL-2001.
XX 18-JAN-2001; 2001WO-US01664.
XX 20-JAN-2000; 2000US-0177109.
PR 14-AUG-2000; 2000US-0225079.
XX (BGMH ) BRIGHAM & WOMENS HOSPITAL INC.
XX Kroll TG, Fletcher JA;
XX WPI; 2001-514487/56.
DR P-PSDB; AAB85796.
XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
PT treating certain tumors or cancers, e.g. follicular carcinoma
XX Claim 21; Page 118; 145pp; English.
XX The invention relates to an oncogene designated PAX8-PPARGammael that
CC contains a PAX8 coding region fused to PPARGammael coding region. The
CC PAX8-PPARGammael polypeptides can be expressed by standard recombinant
CC methodology. A PPARGammael ligand or agent is useful for treating a
CC subject having a disorder characterized by the presence of a PAX8-
CC PPARGammael, where the disorder is cancer, e.g. follicular carcinoma.
CC The PAX8-PPARGammael molecules are also useful for providing nucleotide
CC and amino acid sequences useful for detecting the above disease. The
CC present sequence represents the nucleotide sequence including and
CC surrounding the translocation fusion juncture in the PAX8e7-PPARGammael
CC DNA.
XX Sequence 42 BP; 12 A; 11 C; 12 G; 7 T; 0 other;
XX Query Match 60.0%; Score 25.2; DB 22; Length 42;
XX Best Local Similarity 78.9%; Pred. No. 0.41;
XX Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 5 CCAGGCGCCCTCTCTCAGAAATGACCATGGTTGACACAG 42
DB 5 CTACCCCGCTGGTGGCAGAAATGACCATGGTTGACACAG 42
RESULT 10
AAH76281
ID AAH76281 standard; cDNA; 2334 BP.
XX AAH76281;
XX 29-OCT-2001 (first entry)
XX Human PAX8e7-PPARGammael cDNA sequence.
DE Human PAX8e7-PPARGammael cDNA sequence.
DE PAX8-PPARGammael; oncogene; cytostatic; PAX8; PPARGammael; cancer;
KW follicular carcinoma; PAX8e7-PPARGammael; human; ss.
KW Homo sapiens.
OS Homo sapiens.
XX Key Location/Qualifiers
PH Key Location/Qualifiers
FT CDS 1..2334
FT /*tag= a
XX WO200152789-A2.
XX 26-JUL-2001.
XX 18-JAN-2001; 2001WO-US01664.
XX 20-JAN-2000; 2000US-0177109.
```

```

PR 14-AUG-2000; 2000US-0225079.
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX Kroll TG, Fletcher JA;
XX WPI; 2001-514487/56.
DR P-PSDB; AAB85793.
XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
PT treating certain tumors or cancers, e.g. follicular carcinoma -
XX Claim 1; Page 100-104; 145pp; English.
XX The invention relates to an oncogene designated PAX8-PPARGamma1 that
CC contains a PAX8 coding region fused to PPARGamma1 coding region. The
CC PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant
CC methodology. A PPARGamma1 ligand or agent is useful for treating a
CC subject having a disorder characterized by the presence of a PAX8-
CC PPARGamma1, where the disorder is cancer, e.g. follicular carcinoma.
CC The PAX8-PPARGamma1 molecules are also useful for providing nucleotide
CC and amino acid sequences useful for detecting the above disease. The
CC present sequence represents a human PAX8e7-PPARGamma1 polypeptide
CC encoding cDNA.
XX Sequence 2334 BP; 614 A; 650 C; 573 G; 497 T; 0 other;
SQ
Query Match 60.0%; Score 25.2; DB 22; Length 2334;
Best Local Similarity 78.9%; Pred. No. 1;
Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 CCAGCCCTCTCTCAGAAATGACCATGGTTGACACAG 42
DB 882 CTACCCCGTGTGGCAGAAATGACCATGGTTGACACAG 919

RESULT 11
AAH76286
ID AAH76286 standard; DNA; 42 BP.
XX AC AAH76286;
XX 29-OCT-2001 (first entry)
XX PAX8e9-PPARGamma1 fusion junction DNA sequence.
XX PAX8-PPARGamma1; oncogene; cytostatic; PAX8; PPARGamma1; cancer;
KW follicular carcinoma; PAX8e9-PPARGamma1; human; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 3..41
FT CDS /*tag= a
FT FT
XX WO200152789-A2.
XX 26-JUL-2001.
XX 18-JAN-2001; 2001WO-US01664.
XX 20-JAN-2000; 2000US-0177109.
PR 14-AUG-2000; 2000US-0225079.
XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX Kroll TG, Fletcher JA;
XX WPI; 2001-514487/56.
DR P-PSDB; AAB85798.
XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
PT treating certain tumors or cancers, e.g. follicular carcinoma -

PR 14-AUG-2000; 2000US-0225079.
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX Kroll TG, Fletcher JA;
XX WPI; 2001-514487/56.
DR P-PSDB; AAB85793.
XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
PT treating certain tumors or cancers, e.g. follicular carcinoma -
XX Claim 1; Page 100-104; 145pp; English.
XX The invention relates to an oncogene designated PAX8-PPARGamma1 that
CC contains a PAX8 coding region fused to PPARGamma1 coding region. The
CC PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant
CC methodology. A PPARGamma1 ligand or agent is useful for treating a
CC subject having a disorder characterized by the presence of a PAX8-
CC PPARGamma1, where the disorder is cancer, e.g. follicular carcinoma.
CC The PAX8-PPARGamma1 molecules are also useful for providing nucleotide
CC and amino acid sequences useful for detecting the above disease. The
CC present sequence represents a human PAX8e7-PPARGamma1 polypeptide
CC encoding cDNA.
XX Sequence 2334 BP; 614 A; 650 C; 573 G; 497 T; 0 other;
SQ
Query Match 58.1%; Score 24.4; DB 22; Length 42;
Best Local Similarity 82.4%; Pred. No. 0.9;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0.

QY 9 GGCCTCTCTCTCAGAAATGACCATGGTTGACACAG 42
DB 9 GGCATGCTGTGGCAGAAATGACCATGGTTGACACAG 42

RESULT 12
AAH76294
ID AAH76294 standard; cDNA; 2596 BP.
XX AC AAH76294;
XX 29-OCT-2001 (first entry)
XX Human PAX8e9(-exon 8)-PPARGamma1 cDNA sequence.
XX PAX8-PPARGamma1; oncogene; cytostatic; PAX8; PPARGamma1; cancer;
KW follicular carcinoma; PAX8e9(-exon 8)-PPARGamma1; human; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 161..2596
FT CDS /*tag= a
FT FT
XX WO200152789-A2.
XX 26-JUL-2001.
XX 18-JAN-2001; 2001WO-US01664.
XX 20-JAN-2000; 2000US-0177109.
PR 14-AUG-2000; 2000US-0225079.
XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX Kroll TG, Fletcher JA;
XX WPI; 2001-514487/56.
DR P-PSDB; AAB85801.
XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
PT treating certain tumors or cancers, e.g. follicular carcinoma -
XX Claim 1; Page 127-131; 145pp; English.
XX The invention relates to an oncogene designated PAX8-PPARGamma1 that
CC contains a PAX8 coding region fused to PPARGamma1 coding region. The
CC PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant
CC methodology. A PPARGamma1 ligand or agent is useful for treating a
CC subject having a disorder characterized by the presence of a PAX8-
CC PPARGamma1, where the disorder is cancer, e.g. follicular carcinoma.
CC The PAX8-PPARGamma1 molecules are also useful for providing nucleotide
CC and amino acid sequences useful for detecting the above disease. The

```

CC present sequence represents a human PAX8e9(-exon 8)-PPARGammael
CC polypeptide encoding cDNA.

SQ Sequence 2596 BP; 660 A; 745 C; 672 G; 519 T; 0 other;
Query Match 58.1%; Score 24.4; DB 22; Length 2596;
Best Local Similarity 82.4%; Pred. No. 2.3;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 9 GCCTCTCTCAGAAATGACCATGGTTGACACAG 42
DB 1148 GGATGCTGGCAGAAATGACCATGGTTGACACAG 1181

RESULT 13

AAH76283

ID AAH76283 standard; cDNA; 2625 BP.

XX AAH76283;

AC 19-OCT-2001 (first entry)

XX Human PAX8e9-PPARGammael cDNA sequence.

DE PAX8-PPARGammael; oncogene; cytostatic; PAX8; PPARGammael; cancer;

XX follicular carcinoma; PAX8e9-PPARGammael; human; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

PH CDS 1..2625

FT /*tag= a

XX WO200152789-A2.

XX 26-JUL-2001.

XX 19-JAN-2001; 2001WO-US01664.

XX 20-JAN-2000; 2000US-0177109.

XX 14-AUG-2000; 2000US-0225079.

XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.

XX Kroll TG, Fletcher JA;

XX WPI; 2001-514487/56.

XX P-PSDB; AAB85795.

XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and

XX treating certain tumors or cancers, e.g. follicular carcinoma

XX Claim 1, Page 112-116; 145pp; English.

XX The invention relates to an oncogene designated PAX8-PPARGammael that

XX contains a PAX8 coding region fused to PPARGammael coding region. The

XX PAX8-PPARGammael polypeptides can be expressed by standard recombinant

XX methodology. A PPARGammael ligand or agent is useful for treating a

XX subject having a disorder characterized by the presence of a PAX8-

XX PPARGammael, where the disorder is cancer, e.g. follicular carcinoma.

XX The PAX8-PPARGammael molecules are also useful for providing nucleotide

XX and amino acid sequences useful for detecting the above disease. The

XX present sequence represents a human PAX8e9-PPARGammael polypeptide

XX encoding cDNA.

XX Sequence 2625 BP; 662 A; 761 C; 646 G; 556 T; 0 other;

XX Query Match 58.1%; Score 24.4; DB 22; Length 2625;

XX Best Local Similarity 82.4%; Pred. No. 2.4;

XX Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 9 GCCTCTCTCAGAAATGACCATGGTTGACACAG 42

DB 1177 GGATGCTGGCAGAAATGACCATGGTTGACACAG 1210

RESULT 14

AAAX00271

ID AAAX00271 standard; DNA; 26 BP.

XX AAAX00271,

XX 25-MAR-1999 (first entry)

XX Human PPARG-gamma PCR sense primer F3.

XX Human; Fas; PCR primer; reporter gene; cancer; autoimmune disease;

XX mouse fibroblast L929; human cancer cell HeLa; receptor; PPARG alpha;

XX PPARG gamma; PPARG delta; ss.

XX Synthetic.

XX Homo sapiens.

XX WO9900491-A1.

XX 07-JAN-1999.

XX 23-JUN-1998; 98WO-JP02785.

XX 27-JUN-1997; 97JP-0171440.

XX (ONOH) ONO PHARM CO LTD.

XX Hagiya H, Minami M, Tajima H;

XX WPI; 1999-095737/08.

XX Novel reporter gene DNA-containing plasmid - useful as remedies for

XX cancer or autoimmune diseases, with transformant applicable for

XX detecting ligand of intracellular receptor

XX Example 2; Page 18; 43pp; Japanese.

XX The present invention describes plasmid DNA containing a DNA encoding an

XX amino-acid sequence with a promoter downstream from the Gal 4 protein

XX response sequence, as well as a Fas antigen membrane linking region and

XX function expression region. Also described are: (i) a transformant cell

XX containing both the plasmid DNA and effector protein encoding DNA; (ii)

XX a method by using the transformant cells for screening agonists or

XX antagonists for the intracellular receptor; and (iii) a similar method

XX for screening intracellular receptor agonists or antagonists in mouse

XX fibroblast L929 or human cancer cell HeLa, with the receptor particularly

XX being PPARG alpha, gamma or delta. The plasmid DNA and effector protein

XX encoding DNA are used as active ingredient in drugs to treat cancer or

XX autoimmune diseases, as is the transformant cells. The transformant cells

XX can also be applied in the detection of ligand of an intracellular

XX receptor in a screening method for the intracellular receptor agonists

XX or antagonists, particularly in mouse fibroblast L929 or human cancer

XX cell HeLa with receptor being PPARG alpha, gamma or delta. The present

XX sequence represents a PCR primer used in an example from the present

XX invention.

XX SQ Sequence 26 BP; 11 A; 4 C; 7 G; 4 T; 0 other;

XX Query Match 54.8%; Score 23; DB 20; Length 26;

XX Best Local Similarity 100.0%; Pred. No. 3.1;

XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AGAAATGACCATGGTTGACACAG 42

DB 1 AGAAATGACCATGGTTGACACAG 23

RESULT 15

AAD21022

ID AAD21022 standard; cDNA; 1518 BP.

```

XX AAD21022;
XX AC
XX DT 15-JAN-2002 (first entry)
XX DE Human peroxisome proliferator activated receptor (PPAR) gamma2 CDNA.
XX KW Human; peroxisome proliferator activated receptor gamma2; PPAR gamma2;
XX cytotoxic; antiproliferative; antiviral; cancer; cell proliferation;
XX viral infection; pharmaceutical; thiazolidinedione; ss.
XX OS Homo sapiens.
XX FH
XX Key Location/Qualifiers
XX CDS 1..1518
XX FT /*tag= a
XX FT /product= "Human peroxisome proliferator activated
XX FT receptor (PPAR) gamma2 protein"
XX PN US6294559-B1.
XX PD 25-SEP-2001.
XX PF 03-AUG-1998; 98US-0128142.
XX PR 02-MAY-1996; 96US-016694P.
XX PR 18-APR-1997; 97US-0844007.
XX PA (MERI ) MERCK & CO INC.
XX PI Smith RG;
XX DR WPI; 2001-647265/74.
XX DR P-PSDB; AAE12868.
XX PT Use of thiazolidinedione for treating cancer and viral infections -
XX PS Example 3; Fig 3; 17pp; English.
XX CC The invention relates to compounds and ligands that bind to human
XX peroxisome proliferator activated receptors (PPAR) gamma1 and gamma2.
XX The invention is useful for treating cancer and other disorders including
XX excessive cell proliferation and viral infection. The invention is also
XX directed to the use of PPAR gamma1 and gamma2 to identify compounds that
XX are antiproliferative, antiviral and anticancer agents. The invention
XX also relates to a method of treating cancer using a pharmaceutical
XX composition comprising thiazolidinedione in an amount sufficient to
XX modulate PPAR gamma1 and gamma2 activity. The present CDNA sequence
XX encodes human peroxisome proliferator activated receptor (PPAR) gamma2
XX protein related to the invention.
XX SQ Sequence 1518 BP; 438 A; 363 C; 346 G; 371 T; 0 other;

Query Match 54.8%; Score 23; DB 22; Length 1518;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AGAATGACCATGTTGACACAG 42
Db 81 AGAATGACCATGTTGACACAG 103

```

Search completed: January 21, 2003, 23:09:54
 Job time : 129.667 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 23:03:25 ; Search time 26.3333 Seconds
(without alignments)
489.130 Million cell updates/sec

Title: US-09-765-111a-9
Perfect score: 42
Sequence: 1 cgggcccaggccctctctca.....aatgaccattgtgacacag 42

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

T number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:
1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/1/ina/pctus_COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	61.9	1844	4	US-08-134-557D-1
2	23	54.8	1518	4	US-09-128-142-3
3	23	54.8	1608	3	US-09-484-345-3
4	23	54.8	1679	4	US-09-514-247A-5
5	21.4	51.0	277	3	US-08-917-653-4
6	21.4	51.0	373	3	US-08-917-653-3
7	21.4	51.0	915	4	US-09-724-864-14
8	21.4	51.0	1736	4	US-09-255-392-1
9	21.4	51.0	2005	2	US-08-484-200-1
10	21.4	51.0	2005	2	US-08-477-493-1
11	21.4	51.0	2005	3	US-08-465-375-1
12	21.4	51.0	2005	4	US-09-788-070-1
13	21	50.0	1393	1	US-08-464-148-3
14	21	50.0	1393	1	US-08-385-500-3
15	21	50.0	1393	1	US-08-846-784-3
16	20.2	48.1	1291	2	US-08-887-365-34
17	20.2	48.1	1291	2	US-08-887-365-35
18	20.2	48.1	2204	3	US-09-257-799-47
19	20.2	48.1	2204	3	US-08-920-919A-47
20	20	47.6	2085	2	US-08-668-128B-7
21	20	47.6	2085	2	US-08-905-445-7
22	19.8	47.1	8802	3	US-08-896-449A-1
23	19.8	47.1	8802	3	US-09-132-652-1
24	19.8	47.1	40328	3	US-08-742-185-102
25	19.6	46.7	2889	1	US-08-537-002A-4
26	19.6	46.7	2889	3	US-08-863-010-4
27	19.6	46.7	2889	4	US-08-024-429-4

C 28	19.6	46.7	3600	1	US-08-537-002A-5	Sequence 5, Appli
C 29	19.6	46.7	3600	3	US-08-863-010-5	Sequence 5, Appli
C 30	19.6	46.7	3600	4	US-09-024-429-5	Sequence 5, Appli
C 31	19.4	46.2	1982	4	US-09-221-017B-1068	Sequence 1068, Ap
C 32	19.4	46.2	8257	4	US-09-484-970B-65	Sequence 65, Appli
C 33	19.2	45.7	627	4	US-08-981-030-2	Sequence 2, Appli
C 34	19.2	45.7	627	4	US-08-981-030-12	Sequence 12, Appli
C 35	19.2	45.7	682	4	US-08-981-030-1	Sequence 1, Appli
C 36	19.2	45.7	946	2	US-08-408-095-26	Sequence 26, Appli
C 37	19.2	45.7	1129	4	US-09-227-357-40	Sequence 40, Appli
C 38	19.2	45.7	4403765	4	US-09-103-840A-2	Sequence 2, Appli
C 39	19.2	45.7	4411529	4	US-09-103-840A-1	Sequence 1, Appli
C 40	19	45.2	574	1	US-08-264-534-14	Sequence 14, Appli
C 41	19	45.2	574	1	US-08-083-590A-6	Sequence 6, Appli
C 42	19	45.2	574	1	US-08-468-500-14	Sequence 14, Appli
C 43	19	45.2	574	2	US-08-346-126-14	Sequence 14, Appli
C 44	19	45.2	574	2	US-08-346-128-14	Sequence 14, Appli
C 45	19	45.2	574	3	US-08-532-384-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1

US-08-134-557D-1
; Sequence 1, Application US/08134557D
; Patent No. 6200802
; GENERAL INFORMATION:
; APPLICANT: Greene, Marianne E.
; APPLICANT: Blumberg, Bruce
; TITLE OF INVENTION: Human Peroxisome Proliferator Activated
; TITLE OF INVENTION: Receptor Gamma: Compositions and Methods
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rockey, Milnamow & Katz, Ltd.
; STREET: 2 Prudential Plaza, Suite 4700 180 N. Stetson
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/134.557D
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Katz, Martin L.
; REGISTRATION NUMBER: 25,011
; REFERENCE/DOCKET NUMBER: ARCH:098
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5400
; TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1844 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 179..1606
; US-08-134-557D-1

Query Match 61.9%; Score 26; DB 4; Length 1844;
Best Local Similarity 85.3%; Pred. No. 0.047;
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 GCCCTCTCTCAGAAATGACCATGGTTCACACAG 42

Db 158 GCGTGGCGGAGAAATGACCATGGTTGACACAG 191
|||||
RESULT 2
US-09-128-142-3
; Sequence 3, Application US/09128142
; Patent No. 6294559
; GENERAL INFORMATION:
; APPLICANT: Smith, Roy G.
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENTS ASSOCIATED WITH
; PEROXISOME PROLIFERATOR ACTIVATED RECEPTORS GAMMA1 AND GAMMA2
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jack L. Tribble
; STREET: 126 E. Lincoln Ave., P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Power Macintosh 7500/100
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128,142
; FILING DATE: 03-AUG-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,007
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Tribble, Jack L.
; REGISTRATION NUMBER: 32,633
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-5321
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1518 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-128-142-3
Query Match 54.8%; Score 23; DB 4; Length 1518;
Best Local Similarity 100.0%; Pred. No. 0.88;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 AGAAATGACCATGGTTGACACAG 42
|||||
Db 81 AGAAATGACCATGGTTGACACAG 103
|||||
RESULT 3
US-09-484-345-3
; Sequence 3, Application US/09484345
; Patent No. 6159734
; GENERAL INFORMATION:
; APPLICANT: Robert McKay
; APPLICANT: Alexander H. Borchers
; APPLICANT: Brenda F. Baker
; TITLE OF INVENTION: ANTISENSE MODULATION OF PEROXISOME PROLIFERATOR-ACTIVATED RECEPTOR
; FILE REFERENCE: RTS-0104
; CURRENT APPLICATION NUMBER: US/09/484,345
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 3
; LENGTH: 1608

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (91)...(1608)
US-09-484-345-3
Query Match 54.8%; Score 23; DB 3; Length 1608;
Best Local Similarity 100.0%; Pred. No. 0.89;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 AGAAATGACCATGGTTGACACAG 42
|||||
Db 171 AGAAATGACCATGGTTGACACAG 193
|||||
RESULT 4
US-09-514-247A-5
; Sequence 5, Application US/09514247A
; Patent No. 6365361
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIVAKU CO. LTD.
; APPLICANT: TANIGUCHI, Tomoyasu
; APPLICANT: MIZUKAMI, Junko
; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PP
; FILE REFERENCE: TANIGUCHI=6
; CURRENT APPLICATION NUMBER: US/09/514,247A
; CURRENT FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/JP98/03734
; PRIOR FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: JP231084/1997
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)...(1679)
US-09-514-247A-5
Query Match 54.8%; Score 23; DB 4; Length 1679;
Best Local Similarity 100.0%; Pred. No. 0.9;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 AGAAATGACCATGGTTGACACAG 42
|||||
Db 239 AGAAATGACCATGGTTGACACAG 261
|||||
RESULT 5
US-08-917-653-4
; Sequence 4, Application US/08917653
; Patent No. 6004751
; GENERAL INFORMATION:
; APPLICANT: Rosenfield, Robert L.
; TITLE OF INVENTION: IDENTIFICATION OF ACTIVATORS AND
; TITLE OF INVENTION: INHIBITORS OF SEBUM FORMATION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/917,653
;; FILING DATE: Concurrently Herewith
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Highlander, Steven L.
;; REGISTRATION NUMBER: 37,642
;; REFERENCE/DOCKET NUMBER: ARCD:216
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (512) 418-3000
;; TELEFAX: (512) 474-7577
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 277 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-917-653-4

Query Match 51.0%; Score 21.4; DB 3; Length 277;
Best Local Similarity 95.7%; Pred. No. 2.7;
Ches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 20 AGAAATGACCATGGTTGACACAG 42
Db 82 AGAAATGACCATGGTTGACACAG 104

RESULT 6
US-08-917-653-3
; Sequence 3, Application US/08917653
; Patent No. 6004751
; GENERAL INFORMATION:
; APPLICANT: Rosenfield, Robert L.
; TITLE OF INVENTION: IDENTIFICATION OF ACTIVATORS AND
; TITLE OF INVENTION: INHIBITORS OF SEBUM FORMATION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/917,653
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: ARCD:216
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-917-653-3

Query Match 51.0%; Score 21.4; DB 3; Length 373;
Best Local Similarity 95.7%; Pred. No. 3;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 20 AGAAATGACCATGGTTGACACAG 42

Db 120 AGAAATGACCATGGTTGACACAG 142

RESULT 7
US-09-724-864-14/c
; Sequence 14, Application US/09724864
; Patent No. 6380362
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Murison, James G.
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
; FILE REFERENCE: 11000.1050U1
; CURRENT APPLICATION NUMBER: US/09/724,864
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Mouse
US-09-724-864-14

Query Match 51.0%; Score 21.4; DB 4; Length 915;
Best Local Similarity 71.8%; Pred. No. 3.8; 11; Indels 0; Gaps 0;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGGCCAGGCCCTCTCTCAGAAATGACCATGGTTGACAC 40
Db 569 GGGCCAGGATCTCTGTGAGCAGGGCCCATGGTGCCCC 531

RESULT 8
US-09-255-392-1
; Sequence 1, Application US/09255392
; Patent No. 6214850
; GENERAL INFORMATION:
; APPLICANT: Evans, Ronald M.
; APPLICANT: Forman, Barry M.
; TITLE OF INVENTION: MODULATORS OF PEROXISOME PROLIFERATOR
; TITLE OF INVENTION: ACTIVATED RECEPTOR-GAMMA, AND METHODS FOR THE USE THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/255,392
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,493
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9958
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-1995
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 2005 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 352..1776
US-09-255-392-1

Query Match 51.0%; Score 21.4; DB 4; Length 1796;
Best Local Similarity 80.6%; Pred. No. 4.5;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 12 CTCCTCTCAGAAATGACCATGGTTGACACAG 42
| | | | | | | | | | | | | | | | | | | | | |
Db 334 CCCTTTACTGAAATACCATTGTTGACACAG 364

RESULT 9

US-08-484-200-1
; Sequence 1, Application US/08484200
; Patent No. 5861274
; GENERAL INFORMATION:

; APPLICANT: EVANS, RONALD M.
; APPLICANT: FORMAN, BARRY M.
; APPLICANT: KLEWER, STEVEN A.
; APPLICANT: ONG, ESTELITA S.
; TITLE OF INVENTION: NOVEL MEMBERS OF THE STEROID/THYROID
; TITLE OF INVENTION: SUPERFAMILY AND USES THEREFOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
; STREET: 444 SOUTH FLOWER STREET, SUITE 2000
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90071

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,200
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:
; NAME: REITER, STEPHEN E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9971
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-1995
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2005 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 352..1776
US-08-484-200-1

Query Match 51.0%; Score 21.4; DB 2; Length 2005;
Best Local Similarity 80.6%; Pred. No. 4.7;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 12 CTCCTCTCAGAAATGACCATGGTTGACACAG 42
| | | | | | | | | | | | | | | | | | | | | |
Db 334 CCCTTTACTGAAATACCATTGTTGACACAG 364

RESULT 10

US-08-477-493-1
; Sequence 1, Application US/08477493
; Patent No. 5919442
; GENERAL INFORMATION:

; APPLICANT: EVANS, RONALD M.
; APPLICANT: FORMAN, BARRY M.
; TITLE OF INVENTION: MODULATORS OF PEROXISOME PROLIFERATOR
; TITLE OF INVENTION: ACTIVATED RECEPTOR-GAMMA, AND METHODS FOR THE USE THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,493
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:
; NAME: REITER, STEPHEN E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9958
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-1995
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2005 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 352..1776
US-08-477-493-1

Query Match 51.0%; Score 21.4; DB 2; Length 2005;
Best Local Similarity 80.6%; Pred. No. 4.7;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 12 CTCCTCTCAGAAATGACCATGGTTGACACAG 42
| | | | | | | | | | | | | | | | | | | | | |
Db 334 CCCTTTACTGAAATACCATTGTTGACACAG 364

RESULT 11

US-08-465-375-1
; Sequence 1, Application US/08465375A
; Patent No. 6022897
; GENERAL INFORMATION:

; APPLICANT: EVANS, RONALD M.
; APPLICANT: FORMAN, BARRY M.
; TITLE OF INVENTION: SELECTIVE MODULATORS OF PEROXISOME
; TITLE OF INVENTION: PROLIFERATOR ACTIVATED RECEPTOR-GAMMA, AND METHODS FOR THE
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: SALK1470-1
; CURRENT APPLICATION NUMBER: US/08/465,375A

; CURRENT FILING DATE: 1995-06-05
; EARLIER APPLICATION NUMBER: 08/428,559
; EARLIER FILING DATE: 1995-04-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1
; LENGTH: 2005
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (352)...(1776)
US-08-465-375-1

Query Match 51.0%; Score 21.4; DB 3; Length 2005;
Best Local Similarity 80.6%; Pred. No. 4.7;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 12 CTCTCTCAGAAATGACCATGTTGACACAG 42
Db 334 CCCTTACTGAAATACCATGTTGACACAG 364

RESULT 12
US-09-788-070-1
; hence 1, Application US/09788070
; ent No. 6413994

; GENERAL INFORMATION:
; APPLICANT: EVANS, Ronald
; APPLICANT: FORMAN, Barry
; TITLE OF INVENTION: MODULATORS OF PEROXISOME PROLIFERATOR ACTIVATED RECEPTOR-GAMMA,
; FILE REFERENCE: SALK1480-2
; CURRENT APPLICATION NUMBER: US/09/788,070
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/955,302
; PRIOR FILING DATE: 1999-02-22

; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 1
; LENGTH: 2005

; TYPE: DNA
; ORGANISM: Mus Musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (352)...(1776)
US-09-788-070-1

Query Match 51.0%; Score 21.4; DB 4; Length 2005;
Best Local Similarity 80.6%; Pred. No. 4.7;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 12 CTCTCTCAGAAATGACCATGTTGACACAG 42
L 334 CCCTTACTGAAATACCATGTTGACACAG 364

RESULT 13
US-08-464-148-3
; Sequence 3, Application US/08464148
; Patent No. 5710026

; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US

; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,148
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/385,500
; FILING DATE: 08-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.

; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 13952-21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600

; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1393 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 112..1239

; OTHER INFORMATION: /product= "CYTOPLASMIC

; OTHER INFORMATION: ANTIPROTEINASE-3 PROTEIN"

US-08-464-148-3

Query Match 50.0%; Score 21; DB 1; Length 1393;
Best Local Similarity 73.0%; Pred. No. 6.3;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 5 CCAGGCGCTCTCTCAGAAATGACCATGTTGACACA 41

Db 321 CCAGTGCCTTCTCACTGAATGACAGGCTGGCACA 357

RESULT 14

US-08-385-500-3

; Sequence 3, Application US/08385500

; Patent No. 5712117

; GENERAL INFORMATION:

; APPLICANT: Sprecher, Cindy A.

; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend Kourie and Crew

; STREET: Steuart Street Tower, One Market Plaza

; CITY: San Francisco

; STATE: California

; COUNTRY: US

; ZIP: 94105-1493

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/385,500

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Parmelee, Steven W.

; REGISTRATION NUMBER: 31,990

; REFERENCE/DOCKET NUMBER: 13952-21

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 467-9600

; TELEFAX: (415) 543-5043

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1393 base pairs
; TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 112..1239
OTHER INFORMATION: /product= "CYTOPLASMIC
ANTIPROTEINASE-3 PROTEIN"
US-08-385-500-3

Query Match 50.0%; Score 21; DB 1; Length 1393;
Best Local Similarity 73.0%; Pred. No. 6.3;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 5 CCAGGCCCTCTCTCAGAAATGACCATGCTTGACACA 41
|||||
Db 321 CCAGTCGCTTCTCACTGAAGTGAACAAGGCTGGCACA 357

RESULT 15
US-08-846-784-3
; Sequence 3, Application US/08846784
; Patent No. 5747645
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/846,784
; FILING DATE: 30-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/385,500
; FILING DATE: 08-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 13952-21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1393 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 112..1239
; OTHER INFORMATION: /product= "CYTOPLASMIC
; OTHER INFORMATION: ANTIPROTEINASE-3 PROTEIN"
US-08-846-784-3

Query Match 50.0%; Score 21; DB 1; Length 1393;
Best Local Similarity 73.0%; Pred. No. 6.3;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 5 CCAGGCCCTCTCTCAGAAATGACCATGCTTGACACA 41

Db 321 CCAGTCGCTTCTCACTGAAGTGAACAAGGCTGGCACA 357
|||||
Search completed: January 22, 2003, 00:43:14
Job time : 32.3333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 23:59:20 ; Search time 27 Seconds
(without alignments)
693.573 Million cell updates/sec

Title: US-09-765-111A-9

Perfect score: 42

Sequence: 1 cgggcccggccctctctca.....aatgaccattgtgacacag 42

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 393868 seqs, 222934149 residues

T. number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptcdat/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptcdat/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptcdat/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptcdat/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptcdat/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptcdat/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptcdat/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptcdat/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptcdat/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptcdat/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptcdat/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptcdat/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptcdat/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptcdat/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	42	10	US-09-765-111A-9
2	42	100.0	2523	10	US-09-765-111A-3
3	26	61.9	1811	10	US-09-765-111A-26
4	26	61.9	1811	10	US-09-880-107-2212
5	26	61.9	2260	10	US-09-816-828-8
6	26	61.9	2329	10	US-09-816-828-9
7	25.2	60.0	42	10	US-09-765-111A-7
8	25.2	60.0	2334	10	US-09-765-111A-1
9	24.4	58.1	42	10	US-09-765-111A-11
10	24.4	58.1	2596	10	US-09-765-111A-22
11	24.4	58.1	2625	10	US-09-765-111A-5
12	23	54.8	1608	10	US-09-765-111A-15
13	23	54.8	1679	12	US-10-109-886-5
14	22	52.4	343	9	US-10-046-935-990
15	22	52.4	343	9	US-09-878-178-990
16	22	52.4	611	10	US-09-925-299-109
17	21.6	51.4	537	10	US-09-765-111A-36
18	21.6	51.4	1372	10	US-09-765-111A-13
19	21.6	51.4	2711	10	US-09-765-111A-24

20 21.4 51.0 2005 12 US-10-142-373-1 Sequence 1, Appli
c 21 21.4 51.0 27483 10 US-09-764-877-2928 Sequence 2928, Ap
22 453 10 US-09-867-701-4671 Sequence 4671, Ap
c 23 20.8 49.5 4959 9 US-10-029-217A-1 Sequence 1, Appli
c 24 20.8 49.5 4960 9 US-10-029-217A-30 Sequence 30, Appli
25 20.6 49.0 3560 10 US-09-764-869-2411 Sequence 2411, Ap
26 20.6 49.0 3561 10 US-09-764-869-2410 Sequence 2410, Ap
27 20.4 48.6 2316 10 US-09-799-777-111 Sequence 111, Ap
28 20 47.6 755 10 US-09-925-301-206 Sequence 206, App
c 29 20 47.6 4711 12 US-10-044-090-210 Sequence 210, App
30 20 47.6 1503841 9 US-09-946-807-1 Sequence 1, Appli
31 20 47.6 1503841 10 US-09-795-668-1 Sequence 1, Appli
32 20 47.6 1503841 10 US-09-795-668-1 Sequence 1, Appli
c 33 19.8 47.1 411 10 US-09-880-107-2893 Sequence 2893, Ap
34 19.8 47.1 930 10 US-09-886-055-472 Sequence 472, App
35 19.8 47.1 8802 10 US-09-886-900-1 GENERAL INFORMA
c 36 19.6 46.7 683 10 US-09-867-701-6261 Sequence 6261, Ap
37 19.6 46.7 174424 10 US-09-967-768A-314 Sequence 314, App
38 19.4 46.2 311 10 US-09-864-761-30461 Sequence 30461, A
c 39 19.4 46.2 551 10 US-09-925-300-816 Sequence 816, App
c 40 19.4 46.2 1506 10 US-09-815-242-7203 Sequence 7203, Ap
41 19.4 46.2 5470 10 US-09-764-864-1593 Sequence 1593, Ap
42 19.2 45.7 260 10 US-09-878-574-11856 Sequence 11856, A
43 19.2 45.7 373 10 US-09-878-574-1285 Sequence 1285, Ap
44 19.2 45.7 409 10 US-09-960-352-5129 Sequence 5129, Ap
45 19.2 45.7 445 10 US-09-960-352-7681 Sequence 7681, Ap

ALIGNMENTS

RESULT 1
US-09-765-111A-9
; Sequence 9, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PPARGAMMA NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(41)
US-09-765-111A-9
Query Match 100.0%; Score 42; DB 10; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGCGCCAGCCCTCTCTCAGAAATGACCATTGTTGACACAG 42
Db 1 CGGCGCCAGCCCTCTCTCAGAAATGACCATTGTTGACACAG 42
RESULT 2
US-09-765-111A-3
; Sequence 3, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.

; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PPARGAMMA NUCLEIC ACID MOLECULES
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2523)
US-09-765-111a-3

Query Match 100.0%; Score 42; DB 10; Length 2523;
Best Local Similarity 100.0%; Pred. No. 4.5e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCCAGCCCTCTCAGAAATGACCATGTTGACACAG 42
Db 1067 CGCGCCAGCCCTCTCAGAAATGACCATGTTGACACAG 1108

RESULT 3
US-09-765-111a-26
; Sequence 26, Application US/09765111a
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PPARGAMMA NUCLEIC ACID MOLECULES
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 1811
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (173)...(1609)
US-09-765-111a-26

Query Match 61.9%; Score 25; DB 10; Length 1811;
Best Local Similarity 85.3%; Pred. No. 0.17;
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 GCCTCTCTCAGAAATGACCATGTTGACACAG 42
Db 158 GCGTGGCGCAGAAATGACCATGTTGACACAG 191

RESULT 4
US-09-880-107-2212
; Sequence 2212, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.

; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2212
; LENGTH: 1811
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L40904
US-09-880-107-2212

Query Match 61.9%; Score 26; DB 10; Length 1811;
Best Local Similarity 85.3%; Pred. No. 0.17;
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 GCCTCTCTCAGAAATGACCATGTTGACACAG 42
Db 158 GCGTGGCGCAGAAATGACCATGTTGACACAG 191

RESULT 5
US-09-816-828-8/c
; Sequence 8, Application US/09816828
; Patent No. US20020150898A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhao, Qing A.
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and
; FILE REFERENCE: 791CIP2E
; CURRENT APPLICATION NUMBER: US/09/816,828
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 8
; LENGTH: 2260
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1612)..(2142)
US-09-816-828-8

Query Match 61.9%; Score 26; DB 10; Length 2260;
Best Local Similarity 85.3%; Pred. No. 0.18;
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 GCCTCTCTCAGAAATGACCATGTTGACACAG 42
Db 1654 GCGTGGCGCAGAAATGACCATGTTGACACAG 1621

RESULT 6

US-09-816-828-9/c
; Sequence 9, Application US/09816828
; Patent No. US20020150898A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhao, Qing A.
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: NO. US20020150898A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; IE REFERENCE: 791CIP2E
; CURRENT APPLICATION NUMBER: US/09/816,828
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: pvc_FL_genes Version 2.0
; SEQ ID NO 9
; LENGTH: 2329
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1612)...(2211)
US-09-816-828-9

Query Match 61.9%; Score 26; DB 10; Length 2329;
Best Local Similarity 85.3%; Pred. No. 0.18;
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 GCCTCTCTCAGAAATGACCATGTTGACACAG 42
||| |
Db 1654 GCGTGGCGGAGAAATGACCATGTTGACACAG 1621

RESULT 7

US-09-765-111A-7
; Sequence 7, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: FAX8-PPARGAMMA NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(41)

Query Match 61.9%; Score 26; DB 10; Length 2329;
Best Local Similarity 85.3%; Pred. No. 0.18;
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 GCCTCTCTCAGAAATGACCATGTTGACACAG 42
||| |
Db 1654 GCGTGGCGGAGAAATGACCATGTTGACACAG 1621

US-09-765-111A-7

Query Match 60.0%; Score 25.2; DB 10; Length 42;
Best Local Similarity 78.9%; Pred. No. 0.14;
Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 5 CCAGGCCCTCTCTCAGAAATGACCATGTTGACACAG 42
||| |
Db 5 CTACCCCGTGGTGGCAGAAATGACCATGTTGACACAG 42

RESULT 8

US-09-765-111A-1
; Sequence 1, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: FAX8-PPARGAMMA NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2334
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2334)
US-09-765-111A-1

Query Match 60.0%; Score 25.2; DB 10; Length 2334;
Best Local Similarity 78.9%; Pred. No. 0.38;
Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 CCAGGCCCTCTCTCAGAAATGACCATGTTGACACAG 42
||| |
Db 882 CTACCCCGTGGTGGCAGAAATGACCATGTTGACACAG 919

RESULT 9

US-09-765-111A-11
; Sequence 11, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: FAX8-PPARGAMMA NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(41)
US-09-765-111A-11

Query Match 60.0%; Score 25.2; DB 10; Length 2334;
Best Local Similarity 78.9%; Pred. No. 0.38;
Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 CCAGGCCCTCTCTCAGAAATGACCATGTTGACACAG 42
||| |
Db 882 CTACCCCGTGGTGGCAGAAATGACCATGTTGACACAG 919

```

Query Match      58.1%; Score 24.4; DB 10; Length 2625;
Best Local Similarity 82.4%; Pred. No. 0.84;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      9 GCCTCCCTCTCAGAAATGACCATGGTGCACACAG 42
          ||| | | | | | | | | | | | | | | |
Db     1177 GCATGTGGCAGAAATGACCATGGTGCACACAG 1210

RESULT 12
US-09-765-111A-15
; Sequence 15, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PPARGgamma NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCES: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 1608
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (91)...(1608)
US-09-765-111A-15

```

```

Query Match          54.8%; Score 23; DB 10; Length 1608;
Best local Similarity 100.0%; Pred.No. 2.8;
Matches 23; Conservative 0; Mismatches 0; Gaps 0;

QY      20 AGAAATGACCATGGTTGACACAG 42
      |||||||
Db       171 AGAAATGACCATGGTTGACACAG 193

RESULT 13
US-10-109-886-5
; Sequence 5, Application US/10109886
; Patent No. US20020119499A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO. LTD.
; APPLICANT: TANIGUCHI, Tomoyasu
; APPLICANT: MIZUKAMI, Junko
; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND
; TITLE OF INVENTION: ANTAGONIST TO PPAR
; FILE REFERENCE: TANIGUCHI-6
; CURRENT APPLICATION NUMBER: US/10/109,886
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/514,247
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/JP98/03734
; PRIOR FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: JP231084/1997
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 5
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)..(1679)

```

US-10-109-886-5

Query Match 54.8%; Score 23; DB 12; Length 1679;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AGAATGACCATGGTTGACACAG 42
|||||
Db 239 AGAATGACCATGGTTGACACAG 261

RESULT 14

US-10-046-935-990
; Sequence 990, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; PLICANT: Stolk, John A.
; FILE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 990
; LENGTH: 343
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-046-935-990

Query Match 52.4%; Score 22; DB 9; Length 343;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GAAATGACCATGGTTGACACAG 42
|||||
Db 121 GAAATGACCATGGTTGACACAG 142

RESULT 15

US-09-878-178-990
; Sequence 990, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
; PLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527
; CURRENT APPLICATION NUMBER: US/09/878,178
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 2237
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 990
; LENGTH: 343
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-878-178-990

Query Match 52.4%; Score 22; DB 9; Length 343;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 GAAATGACCATGGTTGACACAG 42
|||||
Db 121 GAAATGACCATGGTTGACACAG 142

Search completed: January 22, 2003, 02:09:58
Job time : 31 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 23:02:45 ; Search time 1003.67 Seconds
(without alignments)
677.725 Million cell updates/sec

Title: US-09-765-111a-9

Perfect score: 42

Sequence:

1 cgggcccaggccctctctca.....aatgaccatggtgacacag 42

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Tc number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_estc.*

9: gb_esti.*

10: gb_est2.*

11: gb_est3.*

12: gb_est4.*

13: gb_est5.*

14: gb_est6.*

15: em_estfun.*

16: em_estom.*

17: gb_gss.*

18: em_gss_hum.*

19: em_gss_inv.*

20: em_gss_pin.*

21: em_gss_vrt.*

22: em_gss_fun.*

23: em_gss_mus.*

24: em_gss_mam.*

25: em_gss_other.*

26: em_gss_pro.*

27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	61.9	903	9	AL543579
2	26	61.9	1141	14	BM922116
3	24	57.1	432	17	BM922116 AGENCOURT
4	24	57.1	1243	14	BM924484
5	23	54.8	730	13	BM924484 AGENCOURT
6	23	54.8	901	13	BM924484 AGENCOURT

7	22.8	54.3	361	10	BE227998
8	22.8	54.3	469	12	BE724296
9	22.8	54.3	633	9	AJ274160
10	22.8	54.3	758	12	BG183368
c 11	22.6	53.8	433	12	BF546300
12	22.6	53.8	453	9	AI556679
13	22.6	53.8	737	14	BQ199668
14	22.6	53.8	860	12	BE900105
c 15	22.4	53.3	309	10	BE685416
c 16	22.4	53.3	330	10	BE684603
c 17	22.4	53.3	361	17	AQ586497
18	22.4	53.3	454	13	BI682587
19	22.4	53.3	500	13	BM288466
c 20	22.4	53.3	501	17	AQ729636
21	22.4	53.3	552	10	AM653703
22	22.4	53.3	578	13	BI535582
23	22.4	53.3	902	13	BI081730
24	22.2	52.9	251	14	T38568
25	22.2	52.9	468	13	BI950167
26	22.2	52.9	557	17	BH105865
c 27	22.2	52.9	673	13	BJ455326
28	22.2	52.9	895	13	BI547966
29	22.2	52.4	308	12	BE865554
c 30	22.2	52.4	364	14	H54473
c 31	22.2	52.4	370	14	H53507
32	22.2	52.4	400	9	AA053612
c 33	22.2	52.4	408	12	BF750332
34	22.2	52.4	445	9	AA010869
35	22.2	52.4	455	9	AA314023
36	22.2	52.4	534	12	BG569798
37	22.2	52.4	540	17	AZ883149
38	22.2	52.4	546	12	BG323933
39	22.2	52.4	551	14	BM790089
40	22.2	52.4	577	14	BM744007
41	22.2	52.4	614	10	AM954835
42	22.2	52.4	618	17	AZ255817
43	22.2	52.4	620	12	BB867033
44	22.2	52.4	621	10	BE373645
45	22.2	52.4	623	17	AA550301

ALIGNMENTS

RESULT 1
AL543579
LOCUS AL543579 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0DI006YI10 5
DEFINITION prime, mRNA sequence.
ACCESSION AL543579
VERSION AL543579.1 GI:12876058
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 903)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. .903
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI006YI10"
/tissue_type="placenta"
/note="Vector: pCWSVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT)-primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9600 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>

BASE COUNT 249 a 228 c 220 g 205 t 1 others

Query Match 61.9%; Score 26; DB 9; Length 903;
Best Local Similarity 85.3%; Pred. No. 17;
Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 GCGCTCTCTCAGAAATGACCATGTTGACACAG 42

|||||

Db 100 GCGGTGGCGGAGAAATGACCATGTTGACACAG 133

RESULT 2

BM922116

LOCUS

DEFINITION AGNCOURT 6707162 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5754203

5', mRNA sequence.

ACCESSION BM922116

VERSION BM922116.1

KEYWORDS

SOURCE

ORGANISM

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1141)

NIH-MGC <http://imgc.ncbi.nih.gov/>.

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM12/91 row: d column: 12

High quality sequence stop: 105.

Location/Qualifiers

1. .1141

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="IMAGE:5754203"

/clone_lib="NIH_MGC_115"

/lab_host="DH10B"

/note="Organ: pooled brain, lung, testis; Vector:

pcmv-sp076; Site 1: NotI; Site 2: EcoRV (destroyed); RNA

source anonymous pool of 6 male brains, age range 23-27; 1

male lung, age 27; and 1 male testis, age 69. Library is

oligo-dT primed and directionally cloned (EcoRV site is

destroyed upon cloning). Average insert size 1.8 kb,

insert size range 1-3 kb. Library is normalized and

enriched for full-length clones and was constructed by C.

Gruber (Invitrogen). Research genetics tracking code

021. Note: this is a NIH MGC Library."

BASE COUNT 298 a 291 c 301 g 259 t 2 others

ORIGIN

Query Match 61.9%; Score 26; DB 14; Length 1141;

Best Local Similarity 85.3%; Pred. No. 18;

Matches 29; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 GCGCTCTCTCAGAAATGACCATGTTGACACAG 42

|||||

Db 72 GCGGTGGCGGAGAAATGACCATGTTGACACAG 105

RESULT 3

LOCUS

DEFINITION

AGNCOURT 6767565 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5760840

5', mRNA sequence.

ACCESSION BM924484

VERSION BM924484.1

KEYWORDS

SOURCE

ORGANISM

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 432)

NIH-MGC <http://imgc.ncbi.nih.gov/>.

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Mahairas GG, Wallace JC, Hood L

High throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 3197 row: L column: 1

Class: BAC ends

High quality sequence stop: 432.

Location/Qualifiers

1. .432

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="IMAGE:5760840"

/clone_lib="CIT Approved Human Genomic Sperm Library D"

/sex="male"

/note="Organ: sperm; Vector: pBelOAC11; BAC Clones in

E-Coli DH10B"

BASE COUNT 130 a 110 c 100 g 92 t

ORIGIN

Query Match 57.1%; Score 24; DB 17; Length 432;

Best Local Similarity 75.0%; Pred. No. 74;

Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0

QY 2 GCGCCAGCCCTCTCTCAGAAATGACCATGTTGACACAC 41

|||||

Db 381 GCGCCAGCCCTCTCTTATGACCAATGCCATGATAACACA 420

RESULT 4

LOCUS

DEFINITION

AGNCOURT 6767565 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5760840

5', mRNA sequence.

ACCESSION BM924484

VERSION BM924484.1

KEYWORDS

SOURCE

ORGANISM

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1243)

NIH-MGC <http://imgc.ncbi.nih.gov/>.

AUTHORS

TITLE

JOURNAL

COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: Agencourt Bioscience Corporation
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1208 row: i column: 01
 High quality sequence start: 95
 High quality sequence stop: 480.

FEATURES
 source
 1. 1243
 /location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5760840"
 /clone_lib="NIH_MGC_116"
 /lab_host="DH10B"
 /note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon, 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH MGC Library."

BASE COUNT 330 a 376 c 272 g 265 t
 ORIGIN

Query Match 57.1%; Score 24; DB 14; Length 1243;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 CAGAAATGACCATGGTTGACACAG 42
 |||
 Db 292 CAGAAATGACCATGGTTGACACAG 315

RESULT 5
 BI0191893/c
 LOCUS 730 bp mRNA linear EST 20-JUN-2001
 DEFINITION 602858760F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:5000102 5', mRNA sequence.
 ACCESSION BI0191893
 VERSION BI0191893.1 GI:14510213
 KEYWORDS EST.
 S' E human.
 ANISM

REFERENCE
 AUTHORS NIH-MGC http://imgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1031 row: g column: 15
 High quality sequence stop: 727.

FEATURES
 source
 1. 730
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5000102"
 /clone_lib="NIH_MGC_10"

/cell_line="MGC36"
 /lab_host="DH10B"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."

BASE COUNT 112 a 241 c 171 g 206 t
 ORIGIN

Query Match 54.8%; Score 23; DB 13; Length 730;
 Best Local Similarity 74.4%; Pred. No. 2.2e+02;
 Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GGGCCAGGCCCTCTCTCAGAAATGACCATGGTTGACAC 40
 |||
 Db 708 GGGCTAGACCCCTCACTCAGAACTAGCAGATGGGAC 670

RESULT 6
 BI820841
 LOCUS 901 bp mRNA linear EST 04-OCT-2001
 DEFINITION 603034063F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175351 5', mRNA sequence.
 ACCESSION BI820841
 VERSION BI820841.1 GI:15932391
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 901)
 AUTHORS NIH-MGC http://imgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11436 row: m column: 16
 High quality sequence stop: 876.

FEATURES
 source
 1. 901
 /location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5175351"
 /clone_lib="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH MGC Library."

BASE COUNT 255 a 214 c 189 g 243 t
 ORIGIN

Query Match 54.8%; Score 23; DB 13; Length 901;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AGAAATGACCATGGTTGACACAG 42
 |||
 Db 207 AGAAATGACCATGGTTGACACAG 229

```

RESULT 7
BE227998
LOCUS      894036F09.x1 C. reinhardtii CC-1690, normalized, Lambda Zap II
DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BE227998
VERSION    BE227998.1 GI:8933237
KEYWORDS   EST.
SOURCE     Chlamydomonas reinhardtii.
           Chlamydomonas reinhardtii
           Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
           Chlamydomonadales; Chlamydomonas.
REFERENCE  1 (bases 1 to 361)
AUTHORS    Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
           McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
TITLE      Analyses of the Chlamydomonas reinhardtii Genome: A Model,
           Unicellular System for Analyzing Gene Function and Regulation in
           Vascular Plants; project phase 2
JOURNAL    Unpublished (2000)
COMMENT     Contract: Elizabeth H. Harris
           DCMB Box 91000
           Duke University
           Durham, NC 27708-1000, USA
           Tel: 919 613 8164
           Fax: 919 613 8177
           Email: chlamy@duke.edu.
FEATURES   Location/Qualifiers
             source
               1..361
               /organism="Chlamydomonas reinhardtii"
               /strain="CC-1690 wild type mt+ 21gr"
               /db_xref="taxon:3055"
               /clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
               II"
             /note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
             XhoI; This library, constructed by John Davies and Jeffrey
             McDermott, combines cDNAs from CC-1690 cells grown to
             mid-log phase in TAP (acetate-containing) medium in the
             light, TAP medium in the dark, HS (minimal) medium in
             ambient levels of CO2 and HS medium bubbled with 5% CO2.
             PolyA mRNA was purified from each sample, pooled and cDNA
             synthesized. The cDNA was directionally cloned into Lambda
             Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
             pBluescript II SK- plasmids were excised from the Lambda
             Zap clones by superinfection with ExAssist (Stratagene)
             phage. The library was normalized using method 4 described
             in Bonaldo et al (1996) Genome Research 6: 791-806."
BASE COUNT 88 a 117 c 86 g 70 t
ORIGIN
Query Match      54.3%; Score 22.8; DB 10; Length 361;
Best Local Similarity 79.4%; Pred. No. 2e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy  2  GGCCAGGCGCTCTCTCAGAAATGACCATGGTT 35
      ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  90  GGGCATGGCTCCCTCAGAAATGACCATGGTT 123

RESULT 8
BE724296
LOCUS      894073G03.x1 C. reinhardtii CC-1690, normalized, Lambda Zap II
DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BE724296
VERSION    BE724296.1 GI:10125592
KEYWORDS   EST.
SOURCE     Chlamydomonas reinhardtii.
           Chlamydomonas reinhardtii
           Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
           Chlamydomonadales; Chlamydomonas.
REFERENCE  1 (bases 1 to 469)
AUTHORS    Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,

```

```

           McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
           Analyses of the Chlamydomonas reinhardtii Genome: A Model,
           Unicellular System for Analyzing Gene Function and Regulation in
           Vascular Plants; project phase 2
JOURNAL    Unpublished (2000)
COMMENT     Contract: Charles Hauser
           DCMB Box 91000
           Duke University
           Durham, NC 27708-1000
           Tel: 919 613 8159
           Fax: 919 613 8177
           Email: chauser@duke.edu.
FEATURES   Location/Qualifiers
             source
               1..469
               /organism="Chlamydomonas reinhardtii"
               /strain="CC-1690 wild type mt+ 21gr"
               /db_xref="taxon:3055"
               /clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
               II"
             /note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
             XhoI; This library, constructed by John Davies and Jeffrey
             McDermott, combines cDNAs from CC-1690 cells grown to
             mid-log phase in TAP (acetate-containing) medium in the
             light, TAP medium in the dark, HS (minimal) medium in
             ambient levels of CO2 and HS medium bubbled with 5% CO2.
             PolyA mRNA was purified from each sample, pooled and cDNA
             synthesized. The cDNA was directionally cloned into Lambda
             Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
             pBluescript II SK- plasmids were excised from the Lambda
             Zap clones by superinfection with ExAssist (Stratagene)
             phage. The library was normalized using method 4 described
             in Bonaldo et al (1996) Genome Research 6: 791-806."
BASE COUNT 105 a 160 c 111 g 93 t
ORIGIN
Query Match      54.3%; Score 22.8; DB 12; Length 469;
Best Local Similarity 79.4%; Pred. No. 2.2e+02;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy  2  GGCCAGGCGCTCTCTCAGAAATGACCATGGTT 35
      ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  90  GGGCATGGCTCCCTCAGAAATGACCATGGTT 123

RESULT 9
AJ274160
LOCUS      AJ274160 Metarhizium anisopliae ARSEF 2575 Metarhizium anisopliae
DEFINITION cDNA clone Ma#1795, mRNA sequence.
ACCESSION  AJ274160
VERSION     AJ274160.1 GI:6433533
KEYWORDS    EST.
SOURCE      Metarhizium anisopliae.
           Metarhizium anisopliae
           Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
           Hypocerales; Clavicipitaceae; mitosporic Clavicipitaceae;
           Metarhizium.
REFERENCE   1 (bases 1 to 633)
AUTHORS    Screen,S.E., Mathur,P. and St. Leger,R.J.
TITLE      EST analysis of the insect pathogenic fungus Metarhizium anisopliae
JOURNAL    Unpublished (1999)
COMMENT     Contact: Screen SE
           Entomology
           University of Maryland
           4112 Plant Sciences Building, College Park, MD 20742, USA.
FEATURES   Location/Qualifiers
             source
               1..633
               /organism="Metarhizium anisopliae"
               /strain="ARSEF 2575"
               /db_xref="taxon:5530"
               /clone_lib="Ma#1795"
               /note="Vector: UniZap; Metarhizium anisopliae was grown on

```



```

ACCESSION      AI556679.1  GI:4489042
VERSION        BQ199668
KEYWORDS       EST.
SOURCE         Norway rat.
ORGANISM       Rattus norvegicus

REFERENCE      1 (bases 1 to 453)
AUTHORS        Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE          Normalization and subtraction: two approaches to facilitate gene
               discovery
JOURNAL        Genome Res. 6 (9), 791-806 (1996)
MEDLINE        97044477
COMMENT        Contact: Soares, MB
               Program for Rat Gene Discovery and Mapping
               University of Iowa
               451 Eckstein Medical Research Building Iowa City, IA 52242, USA
               Tel: 319 335 8250
               Fax: 319 335 9565
               Email: msoares@blue.weeg.uiowa.edu
               The sequence tag present in the cDNA between the NotI site and the
               oligo-dT track served to identify it as a clone from the normalized
               adult spleen library. cDNA Library Preparation: M. Fatima Bonaldo,
               Ph.D. Clone distribution: clones will be available through Research
               Genetics (www.resgen.com) This clone is also available through the
               I.M.A.G.E. Consortium at LNL (info@image.llnl.gov). IMAGE
               ID=1789869 The following repetitive elements were found in this
               cDNA sequence: 1-23. >AT_rich#Low_complexity
               Seq primer: M13 Forward
               POLYA=No.

FEATURES
  source
    1..453
    Location/Qualifiers
      /organism="Rattus norvegicus"
      /strain="Sprague-Dawley"
      /db_xref="taxon:10116"
      /clone="UI-R-C2p-2: c-02-0-UI"
      /clone_lib="UI-R-C2p"
      /dev_stage="adult"
      /lab_hosts="DH10B (Life Technologies)"
      /note="Vector: pT73D-Pac (Pharmacia) with a modified
      polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p
      library is a subtracted library derived from the
      UI-R-C0 library. The UI-R-C0 library consisted of a
      mixture of individually tagged normalized libraries
      constructed from rat placenta, adult lung, brain, liver,
      kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
      embryo. The tag is a string of 3-5 nucleotides present
      between the Not I site and the oligo-dT track which allows
      identification of the library of origin of a clone within
      the mixture. The subtracted library (UI-R-C2p) was
      constructed as follows: PCR amplified cDNA inserts from
      UI-R-C1 clones from which 3' ESTs had been derived was
      used as a driver in a hybridization with the UI-R-C1
      library in the form of single-stranded circles. The
      remaining single-stranded circles (subtracted library) was
      purified by hydroxyapatite column chromatography,
      converted to double-stranded circles and electroporated
      into DH10B bacteria (Life Technologies) to generate the
      UI-R-C2p library. This procedure has been previously
      described (Bonaldo, Lennon and Soares, Genome Research 6:
      791-806, 1996)"
    BASE COUNT  167 a  82 c  81 g  123 t
    ORIGIN
      GGGCCAGGCGCTCTCTCAGAAATGACCATGGTTGAC 38
      |||||
      QY 2
      DB 348 GGGCCAGCTCTCTCGGCTCAGAACTGATCATTACTTGAC 384
      |||||

RESULT 13
LOCUS        BQ199668
DEFINITION   UI-R-D01-clh-k-13-0-UI.s1 UI-R-D01 Rattus norvegicus cDNA clone
               EST.
ACCESSION    BQ199668
VERSION      BQ199668.1  GI:20416133
KEYWORDS     EST.
SOURCE       Norway rat.
ORGANISM     Rattus norvegicus
REFERENCE    1 (bases 1 to 737)
AUTHORS      Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE        Normalization and subtraction: two approaches to facilitate gene
               discovery
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      97044477
COMMENT      Contact: Soares, MB
               Program for Rat Gene Discovery and Mapping
               University of Iowa
               451 Eckstein Medical Research Building Iowa City, IA 52242, USA
               Tel: 319 335 8250
               Fax: 319 335 9565
               Email: msoares@blue.weeg.uiowa.edu
               The sequence contained an oligo-dT track that was present in the
               oligonucleotide that was used to prime the synthesis of first
               strand cDNA and therefore this may represent a bonafide poly A
               tail. The sequence tag present in the cDNA between the NotI site
               and the oligo-dT track served to verify it as a clone from the
               normalized rat cell line R3327-5a library cDNA Library Preparation:
               M.B. Soares Lab Clone distribution: clones will be available
               through Research Genetics (www.resgen.com) The following repetitive
               elements were found in this cDNA sequence: 1-38,
               >Poly_A#Simple_repeat
               Seq primer: M13 Forward
               POLYA=Yes.

FEATURES
  source
    1..737
    Location/Qualifiers
      /organism="Rattus norvegicus"
      /strain="Sprague-Dawley"
      /db_xref="taxon:10116"
      /clones="UI-R-D01-clh-k-13-0-UI"
      /clone_lib="UI-R-D01"
      /dev_stage="adult"
      /lab_hosts="DH10B (Life Technologies)"
      /note="Vector: pT73D-Pac (Pharmacia) with a modified
      polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-D01
      library is a normalized Rat Cell Line R3327-5a library
      (NRSA) constructed in pT73 PAC vector according to the
      procedure described by Bonaldo, Lennon & Soares
      (Normalization and subtraction: Two Approaches to
      Facilitate Gene Discovery, Genome Research 6: 791-806,
      1996). The oligonucleotide used to prime first strand
      synthesis contained the sequence tag GGACTAGATC between
      the Not I cloning site and dT18 stretch. The Rat Cell Line
      R3327-5A was provided by Mary Hendrix of the University of
      Iowa.
      TAG_LIB=UI-R-D01
      TAG_TISSUE=rat cell line R3327-5a
      TAG_SEQ=GGACTAGATC"
    BASE COUNT  232 a  169 c  126 g  209 t  1 others
    ORIGIN
      GGGCCAGGCGCTCTCTCAGAAATGACCATGGTTGAC 38
      |||||
      QY 2
      DB 348 GGGCCAGCTCTCTCGGCTCAGAACTGATCATTACTTGAC 384
      |||||

Query Match      53.8%; Score 22.6; DB 9; Length 453;
Best Local Similarity 75.7%; Pred. No. 2.6e-02;
Matches 28; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 GGGCCAGGCGCTCTCTCAGAAATGACCATGGTTGAC 38
DB 348 GGGCCAGCTCTCTCGGCTCAGAACTGATCATTACTTGAC 384

```


GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 23:02:20 ; Search time 775.333 Seconds
(without alignments)
1576.506 Million cell updates/sec

Title: US-09-765-111a-11

Perfect score: 42

Sequence: 1 ccacgcagcagcgttgca.....aatgaccatggtgacacag 42

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

To, number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_on.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	30	71.4	1766	9	HSPPARGB	X90563 H sapiens m
2	30	71.4	1808	9	HUMPPARG	L40904 Homo sapien
3	30	71.4	1811	6	AX409565	AX409565 Sequence
4	30	71.4	1844	6	ARI39020	ARI39020 Sequence
5	27.2	64.8	14697	1	AB013354	AB013354 Methanosa
C 6	24.8	59.0	155438	2	AP005513	AP005513 Oryza sat
C 7	24.8	59.0	185387	2	AP004768	AP004768 Oryza sat
8	24.2	57.6	1518	9	HSU63415	U63415 Human perox
9	24.2	57.6	1608	6	ARI21467	ARI21467 Sequence
10	24.2	57.6	1608	9	HSU79012	U79012 Human ligan
11	24.2	57.6	1679	6	AR203332	AR203332 Sequence
12	24.2	57.6	1679	9	HUMPPARG	D83233 Homo sapien
13	24	57.1	416	9	AB005521S1	AB005521 Homo sapi
14	24	57.1	135667	9	AC093174	AC093174 Homo sapi
C 15	24	57.1	166043	9	AC090947	AC090947 Homo sapi
C 16	24	57.1	185608	9	AC027126	AC027126 Homo sapi
C 17	24	57.1	200171	2	AC130129	AC130129 Rattus no
18	24	57.1	339681	1	AP003009	AP003009 Mesorhizo
19	23.6	56.2	33	6	AX151002	AX151002 Sequence
20	23.6	56.2	110829	2	AF000440	AF000440 Homo sapi
21	23.6	56.2	120700	9	AP001372	AP001372 Homo sapi
C 22	23.6	56.2	158270	2	AC025956	AC025956 Homo sapi
C 23	23.6	56.2	174175	2	AC011719	AC011719 Homo sapi
24	23.6	56.2	218379	2	AC018959	AC018959 Homo sapi
C 25	23.4	55.7	143135	2	AC094726	AC094726 Rattus no
C 26	23.4	55.7	156301	2	AC112834	AC112834 Rattus no
27	23.4	55.7	194774	2	AC107836	AC107836 Mus muscu
28	23	54.8	2757	9	AK023855	AK023855 Homo sapi
C 29	23	54.8	114815	2	AC119459	AC119459 Rattus no
30	23	54.8	146749	2	AC097171	AC097171 Rattus no
C 31	23	54.8	179937	9	AC016683	AC016683 Homo sapi
32	22.8	54.3	1711	4	EFPPARG1	Y12419 B.taurus mr
33	22.8	54.3	10185	1	AC009474	AB009474 Brucella
34	22.8	54.3	106359	2	AC103443	AC103443 Rattus no
35	22.8	54.3	127680	9	AC100825	AC100825 Homo sapi
C 36	22.8	54.3	143479	9	AC090259	AC090259 Homo sapi
C 37	22.8	54.3	154597	9	AP004295	AP004295 Homo sapi
C 38	22.8	54.3	158389	9	AC011797	AC011797 Homo sapi
C 39	22.8	54.3	162773	9	AC016370	AC016370 Homo sapi
40	22.8	54.3	166725	2	AC111575	AC111575 Rattus no
C 41	22.8	54.3	175664	9	AC027811	AC027811 Homo sapi
42	22.8	54.3	176254	2	AC118414	AC118414 Rattus no
43	22.8	54.3	180575	2	AC126486	AC126486 Rattus no
44	22.6	53.8	1759	9	AF033103	AF033103 Macaca mu
45	22.6	53.8	1765	9	AY048695	AY048695 Macaca fa

ALIGNMENTS

RESULT 1
HSPPARGB
LOCUS
DEFINITION H.sapiens mRNA for peroxisome proliferator-activated receptor gamma.
ACCESSION X90563
VERSION X90563.1 GI:1480099
KEYWORDS peroxisome proliferator-activated receptor gamma.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1766)
Lambe,K.G. and Tugwood,J.D.
A human peroxisome-proliferator-activated receptor-gamma is


```

Db 154 CGCCCGCTGGCCGACGAATGACCATGTTGACACAG 191
|||||
AX409565 1811 bp DNA linear PAT 14-JUN-2002
LOCUS
DEFINITION Sequence 2212 from Patent WO0229103.
ACCESSION AX409565
VERSION AX409565.1 GI:21442270
KEYWORDS
SOURCE
ORGANISM Homo sapiens
REFERENCE
AUTHORS Alvaras, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2212 11-APR-2002;
GENE LOGIC INC (US)
FE ES Location/Qualifiers
source 1..1811
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. L40904"
BASE COUNT 510 a 433 c 422 g 446 t
ORIGIN
Query Match 71.4%; Score 30; DB 6; Length 1811;
Best Local Similarity 86.8%; Pred. No. 0.45;
Matches 33; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 CGCAGCATGTTGGCAGAAATGACCATGTTGACACAG 42
|||||
Db 154 CGCCCGCTGGCCGACGAATGACCATGTTGACACAG 191
|||||
RESULT 4
AR139020 1844 bp DNA linear PAT 16-JUN-2001
LOCUS
DEFINITION Sequence 1 from patent US 6200802.
ACCESSION AR139020
VERSION AR139020.1 GI:14481365
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
TORS Unclassified.
LE Greene, M.E. and Blumberg, B.
Human peroxisome proliferator activated receptor gamma:
compositions and methods
JOURNAL Patent: US 6200802-A 1 13-MAR-2001;
FEATURES
Location/Qualifiers
source 1..1844
/organism="unknown"
BASE COUNT 543 a 433 c 421 g 447 t
ORIGIN
Query Match 71.4%; Score 30; DB 6; Length 1844;
Best Local Similarity 86.8%; Pred. No. 0.45;
Matches 33; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 CGCAGCATGTTGGCAGAAATGACCATGTTGACACAG 42
|||||
Db 154 CGCCCGCTGGCCGACGAATGACCATGTTGACACAG 191
|||||
RESULT 5
AE013354 14697 bp DNA linear BCT 17-MAY-2002
LOCUS
DEFINITION Methanosarcina mazei strain Goel, section 136 of 379 of the
complete genome.
ACCESSION AE013354 AE008384

```

```

AB013354.1 GI:20905689
Methanosarcina mazei Goel.
Methanosarcina mazei Goel.
Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
Methanosarcinaceae; Methanosarcina.
1 (bases 1 to 14697)
Deppenmeier U., Johann, A., Hartsch, T., Merkl, R., Schmitz, R.A.,
Martinez-Arias, R., Henne, A., Wier, A., Baeumer, S., Jacobi, C.,
Bruggemann, H., Lienard, T., Christmann, A., Boemcke, M., Steckel, S.,
Bhattacharyya, A., Lykidis, A., Overbeek, R., Klenk, H.-P.,
Gunsalus, R.P., Fritz, H.-J. and Gottschalk, G.
The genome of Methanosarcina mazei: Evidence for lateral gene
transfer between Bacteria and Archaea
J. Mol. Microbiol. Biotechnol. (2002) In press
2 (bases 1 to 14697)
Deppenmeier, U., Johann, A., Hartsch, T., Merkl, R., Schmitz, R.A.,
Martinez-Arias, R., Henne, A., Wier, A., Baeumer, S., Jacobi, C.,
Bruggemann, H., Lienard, T., Christmann, A., Boemcke, M., Steckel, S.,
Bhattacharyya, A., Lykidis, A., Overbeek, R., Klenk, H.-P.,
Gunsalus, R.P., Fritz, H.-J. and Gottschalk, G.
Direct Submission
Submitted (23-AUG-2001) Goettingen Genomics Laboratory, Institute
for Microbiology and Genetics, Grisebachstrasse 8, Goettingen
37077, Germany
Location/Qualifiers
1..14697
/organism="Methanosarcina mazei Goel"
/db_xref="taxon:192952"
complement (444..1532)
/gene="MM1233"
complement (444..1532)
/gene="MM1232"
/codon_start=1
/transl_table=11
/product="conserved protein"
/protein_id="AA030928.1"
/db_xref="GI:20905690"
/translation="MSSGISLPLKTEHSARDRLYNFLNAPFLIGSIIVARDLSKLV
AVIGTALIFPIIILRFACHTSYARSKGVCKMMFPGVKIFKARPGSHSPFEKA
MTVIGTLPMPFLPVMVIRLDLLGAYIVAWALFLLTARYACVRCINFECPVNRVS
GVKEPKRDPGNIEVSKBEFWAAGSCTSLRVKHEFKDFLYWNEVTTILFSAQ
LAIGISTGMLLAYIFIAFFHGLEORFCTCHPCYVGEKCLRMNWMGPHFKP
RPHPKPDAITSIGFWLILFPPLPFLLEIPFLGLTYLSVSISVFLITINWYECRCI
YFCEPFRVPAEVRKEFEKKDFERMILK"
complement (1668..2486)
/gene="MM1233"
complement (1668..2486)
/gene="MM1233"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA030929.1"
/db_xref="GI:20905691"
/translation="MNTAKESKLIKYKPGKEYIATQCTLSSTHTTITYKGLKEGF
FLKDLAPGLIKIFGKSFSAVCEIPNSLPAIIFPIVSPFSAFPLSVANYTSEGLI
LAILDSGVITTAPEKLQIYIAEYIRNTVCNCKGFIKREKPKIKETISPTIY
RIQITRYWSCRSCGYMVRSESEGVTKKGTWKVSDLAQVCKRCCKGTGTIYKYP
DIRVSEKETMERRYRCKLCNVEDIAATEFFNVPSDGGVSAIENWIDYEDYFG"
2800..3186
/gene="MM1234"
2800..3186
/gene="MM1234"
/codon_start=1
/transl_table=11
/product="conserved protein"
/protein_id="AA030930.1"
/db_xref="GI:20905692"
/translation="NGRKYQSVLGFDWNPYADKTYVFTGQKEPLRKNRNVFVSGN
VREFTQLKENTKDOIWLWGSQVLIKAFLLEDLVDMDIVFVPIILGGIPLFDRIK
EIKLMTNTERYESGLLRTEYNLKP"
3262..4182
gene
CDS

```

```

/ gene="mmr"
/ note="MM1235"
3262. .4182
/ gene="mmr"
/ codon_start=1
/ transl_table=11
/ product="5-methylcytosine-specific restriction enzyme"
/ protein_id="AAM30931.1"
/ db_xref="GI:20905693"
/ translation="MAIPDYQAVMLPLLVKAGGKEHRIQDAIEKLAEEFELSEBERK
ELLPSCQOAI FKNRIGWANTVLYKAGLIGSRKKGVSFTIRGLDVLRLKPPSDVSFL
KRYKEPNEYKQOVLIDKDPVPEAEKGLDPRESLEYOKLNEIYSELISITK
KCSPAFESLVVDVITKMGYGGSDVAGKAGKSHDGDGIDGLIKEDRLGLDVIYIOAK
RWGTVSKPEIQFAGALIGKAKKGVFITTSFSKBALEYADFTGNIVLVDGEMLAR
LMIEYNGVSRVSKYEVKMDTYFEDGVI"
/ gene="MM1236"
4612. .5475
/ gene="MM1236"
4612. .5475
/ EC number="3.4.24.-"
/ codon_start=1
/ transl_table=11
/ product="protease HtpX"
/ protein_id="AAM30932.1"
/ db_xref="GI:20905694"
/ translation="MKNMLKTVLLAALTGLLVIGDYPGGTGMIAFLPAVLNFG
SYWSDKIVLRYAKESPAEENLHRIYDGLVWKAGIPKPYIVOSGNPAFATG
ADPKAAVAATTGILELISYMEGVLAHELHVKNRDPLISAATLAGVVTMLAHW
AOWAIPFGGGRDDGGGIVGILAMAI VAPIAATLIQIAI SRREFPADEGARIS
RKPVALADALEKLEYGNSHYRARSVDQAKESSAHMFIWNLKGGVQSLRTHPYTD
ERVRRLAMKF"
6011. .6664
/ gene="MM1237"
6011. .6664
/ gene="MM1237"
/ EC number="2.1.1.45"
/ codon_start=1
/ transl_table=11
/ product="Thymidylate synthase"
/ protein_id="AAM30933.1"
/ db_xref="GI:20905695"
/ translation="MENKFEITRIIRAKNISDAWYRGLNIWNHGRVITDERGSOIKE
FMLELVENPVYDRIADPTANREERLEEVAKOLISGENIDPEYTYGRLRWNGEV
DQLEYIEKKSPTSRSRATAVTWIPVDTKVNEVFCMLDDFKLRDEKXHLITLRS
HDFGGAPANLYGLSKLLEYVAEKVGEVPGVITTVSISARVYDHDWDWVENIVKGIN"
6768. .8060
/ gene="MM1238"
6768. .8060
/ gene="MM1238"
/ EC number="2.5.1.19"
/ codon_start=1
/ transl_table=11
/ product="3-phosphoshikimate 1-carboxyvinyltransferase"
/ protein_id="AAM30934.1"
/ db_xref="GI:20905696"
/ translations="MRASISKSISKGVFAPPKSVTHRAITLAALSKEISIIHRPLLS
ADTLATIRASEMFGAARREKENLIIOGSKPGIPDDVIDAANSGLTIRLMTWIAGL
TDGTVLTGDSSELTFRPNGLLEVLNRLGAKACSTRGNERAPIVVKYKIGSEVEISG
SISQIFT SALLIACPLAENSTTSLIIGKLSRPYVDVTIEMGLAGVRIHTDDNNGTK
FIIPGOKYDLKQYTVPGDFSSAYLLAAANLEGEITVKNLFPFSKQGDVITDILK
QMADITWDEAGITVVRGRLKAITFDAGSTPDLPVPTVAVLASVARESTRIENAEH
VRYKETORLHALATEUPKMGVSLKEEMDSLTITGGTLEGAEVGHGWDHRIWMSLAIG
MVAQNTITVDTTESYSISYDPDFPKMRNLGAKVKEIPEE"
8193. .9173
/ gene="MM1239"
8193. .9173
/ gene="MM1239"
/ codon_start=1
/ transl_table=11
/ product="methyltransferase"
/ protein_id="AAM30935.1"
/ db_xref="GI:20905697"
/ translation="MEIKRCQGDCKIKPVSEVLKDIELFYKPCRDCCKTEIKFSPLA

```

```

EQNLDEIDNRGNCKGRQLDIWAHVLKVMIDEGINKKANLRACVPLITPGYL
TDYFPLPENSLVILSSMNEKCAERIVKEVPEVKVLKGDARTKVIKDSDSPHIY
ELLAGCDLRCDIIOTPYGALGIYKYQHEIIEPPOVHSFKTEILEKALKOYNNPSVLD
CTCGPSGLITCLKAGARKVFNDDIWHPALETTLINLEANGFPVKFSGSEKELIASGN
NFEVSTDIRELVNYLDEKVDICIVDTFFPQVDTAKIEAAGKLKGVVVI"
complement (9519. .11231)
/ gene="MM1240"
complement (9519. .11231)
/ gene="MM1240"
/ EC number="1.8.-.-"
/ codon_start=1
/ transl_table=11
/ product="Methyl-coenzyme M reductase, alpha subunit"
/ protein_id="AAM30936.1"
/ db_xref="GI:20905698"
/ translation="MAADIFSKFKSKMEVKFTQEGYGSNQAGGDTGTAKFLRLGPE
QDARKAEMLKAGKETAEKGGIAFYNPWHSAGPLAGORAITPVTLSGTDIVAEPPDLHY
VNNAMQOMWDDIRCTCIVGLDMAHETLEKRLGKEVTPETINNHLEVLNHPMGAAYV
QEMVETHPALVDDCIVRVFTGDDLEADEIDKQIVININKMPSSEQAQAQIKASGKIT
WQAIHPTIVSRITDGAQTSRAAMQIGMSFISAYAMCAGEAAVADLSFAAKHAAL
MGEMLPARRARGNEPGLSFEHLADIITQSRVSDPAPAKIALEVVGAGMLYQIWI
SYMGGVGTQYATAAATDDILNNYTDVYDINDKYNGAANLGTDNKVKATLDVVKU
IATESTLYGITYEYKFTALEDBHFGGSRATVLAASGVACALATGNANAGLSGWLS
MYVHEKAGSLGAFGFDLQDCGATNVLISQDDEGLPDELPGPNYVRYNVMVHGQGY
AGTAQAHSRGDAFTVNLKVCFAADDMLMFFNAEPREFFRGARIEFFVPAGERSLI
IPAK"
complement (11242. .11997)
/ gene="MM1241"
complement (11242. .11997)
/ gene="MM1241"
/ EC number="1.8.-.-"
/ codon_start=1
/ transl_table=11
/ product="Methyl-coenzyme M reductase, gamma subunit"
/ protein_id="AAM30937.1"
/ db_xref="GI:20905699"
/ translation="MHEMAYESQYVPGATSVGNRBNKMSGLKLEISDEDLTAVL
GHRAGSDYPSHTPPLAEMGEPACSEAVAPTPGAAGDRVRYVQPADSYNAPATP
YFRSYFAAINRGVDFGLTSGRQIVARERDEMEQCAKQVOMETMTDHALAGMGATVH
GYSVRLQEDGVFMFLDRRLREGSTIINDKQVAIFLDRKVDLGRKPMSEEAARFTTI
YRVNVAFRDAAEVVEMVQRIFDLRTKYGFQPK"
complement (11990. .12610)
/ gene="MM1242"
complement (11990. .12610)
/ gene="MM1242"
/ codon_start=1
/ transl_table=11
/ product="Methyl-coenzyme M reductase operon protein C"
/ protein_id="AAM30938.1"
/ db_xref="GI:20905700"

Query Match 64.8%; Score 27.2; DB 1; Length 14697;
Best Local Similarity 80.0%; Pred. No. 5.8;
Matches 32; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CCATCCAGCGCATGTCGCAAGTACCATGCTGTCACAC 40
DB 7933 CTATAGCAGCATGTTTCAGGAACACGATAGTTCACAC 7972

RESULT 6
AP005513/c 155438 bp DNA linear HTG 11-JUL-2002
LOCUS Oryza sativa (japonica cultivar-group) chromosome 2 clone
DEFINITION OSUNBa0014E22, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION AP005513
VERSION AP005513.1 GI:21728115
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,
clone:OSUNBa0014E22.
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```


Query Match 57.6%; Score 24.2; DB 9; Length 1518;
 Best Local Similarity 78.4%; Pred. No. 83;
 Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 6 GCAGGCATGGTGGCAGAAATGACCATGTTGACACAG 42
 ||| |||
 DB 67 GCAACATATCACAGAAATGACCATGTTGACACAG 103
 ||| |||

RESULT 9
 LOCUS AR121467 1608 bp DNA linear PAT 16-MAY-2001
 DEFINITION Sequence 3 from patent US 6159734.
 ACCESSION AR121467
 VERSION AR121467.1 GI:14105043
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1608)
 AUTHORS McKay, R., Borchers, A.H. and Baker, B.F.
 TITLE Antisense modulation of peroxisome proliferator-activated receptor
 gamma expression
 JOURNAL Patent: US 6159734-A 3 12-DEC-2000;
 FEATURES Location/Qualifiers
 source 1..1608
 /organism="unknown"

BASE COUNT 462 a 380 c 361 g 405 t
 ORIGIN

Query Match 57.6%; Score 24.2; DB 6; Length 1608;
 Best Local Similarity 78.4%; Pred. No. 83;
 Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 6 GCAGGCATGGTGGCAGAAATGACCATGTTGACACAG 42
 ||| |||
 DB 157 GCAACATATCACAGAAATGACCATGTTGACACAG 193
 ||| |||

RESULT 10
 LOCUS HSU79012 1608 bp mRNA linear PRI 03-MAY-1997
 DEFINITION Human ligand activated transcription factor PPARGamma2 mRNA,
 complete cds.
 ACCESSION U79012
 VERSION U79012.1 GI:1711116
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 REFERENCE 1 (bases 1 to 1608)
 AUTHORS Mukherjee, R., Jow, L., Croston, G.E. and Paterniti, J.R. Jr.
 TITLE Identification, characterization, and tissue distribution of human
 peroxisome proliferator-activated receptor (PPAR) isoforms
 PPARGamma2 versus PPARGamma1 and activation with retinoid X
 receptor agonists and antagonists
 JOURNAL J. Biol. Chem. 272 (12), 8071-8076 (1997)
 MEDLINE 97218249
 PUBMED 9065481
 REFERENCE 2 (bases 1 to 1608)
 AUTHORS Mukherjee, R., Jow, L., Croston, G.E. and Paterniti, J.R. Jr.
 TITLE Direct Submission
 JOURNAL Submitted (20-NOV-1996) Cardiovascular Research, Ligand
 Pharmaceuticals, Inc., 9393 Towne Centre Drive, San Diego, CA
 92121, USA
 FEATURES Location/Qualifiers
 source 1..1608
 /organism="Homo sapiens"
 /db xref="taxon:9606"
 /tissue_type="heart"
 91..1608
 /note="nuclear receptor"

CDS

/codon start=1
 /product="ligand activated transcription factor
 PPARGamma2"
 /protein_id="AAC51248.1"
 /db xref="GI:1711117"
 /translation="MGELIGSDIPDSDSFTDTLSANISQEMTMDTEMFWPTNFG
 ISSVDSVMEDSHSFDIKPTFTVDSSISTPHYEDIPRTDPVADYKDYDLQLEY
 OSAIKVEPASPPYSEKTLQYNKPEPNSLMAIECRVCGKAGPHYGVHACEGCK
 GFFRTIRLKIYDRCDLNCRIHKSRNKOYCRFOKCLAVGSHNAILRFRMPQAEK
 EKLARISSIDIDOLNPESADILRALAKHLVDYVKSPPDLTKAKARAILTGKTKDSFV
 IYDMSLMWGEDKIKFKHITPLQESKEVAIRFQCQFRSVEAVQETETAKSIPGF
 VNLDNDQVLLKYGVEHIIITMLASLNKDGVLISGGQGFREFLSKRPKPGDFM
 EPKFEFAVFNALDELDDLAIFIAVILISGDRPGLLNKPKIEDIQDNLQALELQK
 LNHPESSQLFAKLQKMTDLRQIVTEHVQLLQVIKKTETDMSLHPLLQEIYKDYK"

BASE COUNT 462 a 380 c 361 g 405 t
 ORIGIN

Query Match 57.6%; Score 24.2; DB 9; Length 1608;
 Best Local Similarity 78.4%; Pred. No. 83;
 Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 6 GCAGGCATGGTGGCAGAAATGACCATGTTGACACAG 42
 ||| |||
 DB 157 GCAACATATCACAGAAATGACCATGTTGACACAG 193
 ||| |||

RESULT 11
 LOCUS AR203332 1679 bp DNA linear PAT 20-JUN-2002
 DEFINITION Sequence 5 from patent US 6365361.
 ACCESSION AR203332
 VERSION AR203332.1 GI:21499694
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1679)
 AUTHORS Taniguchi, T. and Mizukami, J.
 TITLE Method for identifying or screening agonist and antagonist to PPAR
 JOURNAL Patent: US 6365361 A 5 02-APR-2002;
 FEATURES Location/Qualifiers
 source 1..1679
 /organism="unknown"

BASE COUNT 477 a 405 c 377 g 420 t
 ORIGIN

Query Match 57.6%; Score 24.2; DB 6; Length 1679;
 Best Local Similarity 78.4%; Pred. No. 83;
 Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 6 GCAGGCATGGTGGCAGAAATGACCATGTTGACACAG 42
 ||| |||
 DB 225 GCAACATATCACAGAAATGACCATGTTGACACAG 261
 ||| |||

RESULT 12
 LOCUS HUMPPARG 1679 bp mRNA linear PRI 06-FEB-1999
 DEFINITION Homo sapiens mRNA for PPAR gamma2, complete cds.
 ACCESSION D83233 D83136
 VERSION D83233.1 GI:1199768
 KEYWORDS PPAR gamma2.
 SOURCE Homo sapiens adipose cDNA to mRNA.
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 1679)
 AUTHORS Kato, S.
 TITLE PPAR gamma2
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1679)
 AUTHORS Kato, S.
 TITLE Direct Submission

JOURNAL Submitted (29-JAN-1996) Shigeaki Kato, Tokyo University of Agriculture, Department of Agricultural Chemistry, Sakuragaoka 1-1-1, Setagaya-ku, Tokyo 156, Japan (E-mail: ktake@u-tokyo.ac.jp, Tel: 03-5477-2318, Fax: 03-5477-2619)

COMMENT On or before Mar 17, 1999 this sequence version replaced

gi:2077947, gi:1197807.

D83136: Submitted (18-Jan-1996).

FEATURES
source
1..1679
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="adipose"
159..1679
/codon_start=1
/product="ppar gamma2"
/protein_id="BAA18949.1"
/db_xref="GI:1945622"
/translation="MGETILGSDIPDSFTLTLSANISQEMTNVDETAFTWTFNG
ISSVLSMEDSHSHFDIKPTFTVDFSSISTPHYEDIPFTDPVADYKYDKLOEY
QSAIKVEPAPSYSEKTLQYNKPHBPSNSLMAIECRVCGDKASGHYGVHACEGCK
GFFRTIRLKLIVDRCLNCRHKKSKQYCRFOKCLAVGMSHNAIRFQRTAQACK
EKLLAETSSDIQDLNPGSADLRQALHLYDSVYKSFPLTKAKARAILTGRTDKSPF
VYDMSLAWGDKIKXKHITPELOESKEVAIRIFQGCOPRSVEAIOETIYAKSIPG
FVNLDNDQVTLKGVHEIITVLMASLMKOGVLISGQGFNFRELFKSRKPFQGF
MEKFEPAVKFNALEDSDIAIFIAVILISGDRPGLNVKFIQDNLQALQLEQL
KLNHPSSQLFALKLQKMTDLRQIVTEHVQLVQIKTKTDTMSLPLLEIYKDY"

BASE COUNT 477 a 405 c 377 g 420 t

ORIGIN

Query Match 57.6%; Score 24.2; DB 9; Length 1679;
Best Local Similarity 78.4%; Pred. No. 83;
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 6 GCAGGCGTGGTGGCAGAAATGACCATGTTGACACAG 42
|||||
Db 225 GCACATATCACAGAAATGACCATGTTGACACAG 261
|||||

RESULT 13
AB005521S1 416 bp DNA linear PRI 14-APR-2000

LOCUS Homo sapiens ppar gamma gene for peroxisome proliferator
activated-receptor gamma, exon 1.

DEFINITION AB005521

ACCESSION AB005521.1 GI:2605490

VERSION ppar gamma; ppar gamma common exon1; peroxisome proliferator
activated-receptor gamma.

KEYWORDS 1 of 6

SEGMENT S 'E

ANIM

REFERENCE

AUTHORS Okazawa, H., Mori, H., Tamori, Y., Araki, S., Niki, T., Masugi, J.,
Kawanishi, M., Kubota, T., Sinoda, H. and Kasuga, M.

TITLE No coding mutations are detected in the peroxisome
proliferator-activated receptor- gene in Japanese patients with
lipotrophic diabetes

JOURNAL Diabetes (1997) In press

REFERENCE 2 (bases 1 to 416)

AUTHORS Okazawa, H.

TITLE Direct Submission

JOURNAL Submitted (03-JUL-1997) Hideki Okazawa, Kobe University School of
Medicine, 2nd Department of Internal Medicine; 7-5-1 Kusunoki-cho
Chuo-ku, Kobe 650, Japan (E-mail: okazawa@med.kobe-u.ac.jp,
Tel: 81-78-341-7451, Fax: 81-78-382-2080)

FEATURES
source
1..416
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="placenta"
115..340
/gene="ppar gamma"

exon

/product="peroxisome proliferator activated-receptor
gamma"
/note="ppar gamma common exon1"
/number=1

BASE COUNT 113 a 105 c 67 g 131 t

ORIGIN

Query Match 57.1%; Score 24; DB 9; Length 416;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 CAGAAATGACCATGTTGACACAG 42
|||||

Db 110 CAGAAATGACCATGTTGACACAG 133
|||||

RESULT 14
AC0931174 135667 bp DNA linear PRI 13-AUG-2001

LOCUS Homo sapiens chromosome 3 clone RP11-167M22 map 3p, complete
sequence.

DEFINITION AC0931174

ACCESSION AC0931174.1 GI:15148930

VERSION HTG.

KEYWORDS Homo sapiens.

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 135667)

Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H.,
Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z.,
He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F.,
Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y.,
Li, W., Li, W., Li, Y., Luo, J., Luo, J., Qi, Q., Qi, X., Song, L.,
Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J.,
Wang, J., Wang, J., Wang, L., Wang, R., Wang, X., Wang, X., Wang, X.,
Wang, Y., Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X.,
Yu, P., Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M.,
Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,
Yu, J. and Yang, H.

Chromosome 3p genomic sequence

Unpublished

2 (bases 1 to 135667)

Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J., Ding, H.,
Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D., Guo, Z.,
He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C., Li, F.,
Li, G., Li, J., Li, L., Li, S., Li, T., Liu, Y., Liu, N., Liu, B., Liu, Y.,
Li, W., Li, W., Li, Y., Luo, J., Luo, J., Qi, Q., Qi, X., Song, L.,
Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J.,
Wang, J., Wang, L., Wang, R., Wang, X., Wang, X., Wang, X., Wang, Y.,
Wu, D., Wu, Q., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B.,
Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M., Zhang, M.,
Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,
Yu, J. and Yang, H.

Direct Submission

Submitted (13-AUG-2001) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P. R. China

Center: Beijing Center

Center code: Beijing

Website: http://hgsc.igtp.ac.cn

http://www.genomics.org.cn

Contact: hgsc@igtp.ac.cn

Center project name: \$ project

Center clone name: RP11-167M22

Summary Statistics

Sequencing vector: pUC18; 100% of reads

Chemistry: Dye-terminator; ET 55% of reads

Chemistry: Dye-terminator Big Dye; 45% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 586 bases at least Q40

Consensus quality: 919 bases at least Q30
 Consensus quality: 1154 bases at least Q20
 Insert size: 1198; sum-of-contigs
 Quality coverage: 1.48x in Q20 bases; sum-of-contigs

FEATURES

Location/Qualifiers
 1. 115667
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3"
 /map="3p"
 /clone="RP11-167M22"

BASE COUNT 41244 a 25375 c 26554 g 42494 t

ORIGIN

Query Match 57.1%; Score 24; DB 9; Length 135667;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 CAGAAATGACCATGTTGACACAG 42

Db 109924 CAGAAATGACCATGTTGACACAG 109947

RESULT 15

AC090947/c

LOCUS

DEFINITION Homo sapiens chromosome 3 clone RP11-30G23 map 3p, complete

AC090947 166043 bp DNA linear PRI 20-MAR-2001

AC090947 AC016333

VERSION AC090947.1 GI:13384351

KEYWORDS HTG.

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

REFERENCE 1 (bases 1 to 166043)

1 (bases 1 to 166043)

Wu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J.,

Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D.,

Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C.,

Li, F., Li, G., Li, J., Li, S., Li, T., Liu, Y., Liu, N., Liu, B.,

Liu, Y., Li, W., Li, W., Li, Y., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L.,

Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J.,

Wang, J., Wang, J., Wang, L., Wang, R., Wang, X., Wang, X.,

Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B.,

Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M.,

Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,

Yu, J. and Yang, H.

Chromosome 3p genomic sequence

Unpublished

2 (bases 1 to 166043)

Wu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J.,

Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D.,

Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C.,

Li, F., Li, G., Li, J., Li, S., Li, T., Liu, Y., Liu, N., Liu, B.,

Liu, Y., Li, W., Li, W., Li, Y., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L.,

Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J.,

Wang, J., Wang, J., Wang, L., Wang, R., Wang, X., Wang, X.,

Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B.,

Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M.,

Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,

Yu, J. and Yang, H.

Direct Submission

Submitted (20-MAR-2001) Human Genomic Center, Institute of

Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing

100101, P. R. China

On Mar 20, 2001 this sequence version replaced gi:8072582.

-----Genome Center

Center: Beijing Center

Center code: Beijing

Website: <http://hgsc.igtp.ac.cn>

<http://www.genomics.org.cn>

Contact: hgc@igtp.ac.cn

COMMENT

----- Project Information
 Center project name: 1% project
 Center clone name: RP11-30G23
 ----- Summary Statistics
 Sequencing vector: pUC18; 100% of reads
 Chemistry: Dye-terminator; ET 55% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 499 bases at least Q40
 Consensus quality: 605 bases at least Q30
 Consensus quality: 674 bases at least Q20
 Insert size: 692; sum-of-contigs
 Quality coverage: 2.80x in Q20 bases; sum-of-contigs

FEATURES

source

1. 166043
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3"
 /map="3p"
 /clone="RP11-30G23"

BASE COUNT 48570 a 35113 c 34446 g 47914 t

ORIGIN

Query Match 57.1%; Score 24; DB 9; Length 166043;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 CAGAAATGACCATGTTGACACAG 42

Db 117076 CAGAAATGACCATGTTGACACAG 117053

Search completed: January 21, 2003, 23:59:16
 Job time : 898.333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 23:00:10 ; Search time 127.667 Seconds
(without alignments)
740.866 Million cell updates/sec

Title: US-09-765-111a-11

Perfect score: 42

Sequence:

1 ccacgcgacgatgtggca.....aattgaccatggtgacacag 42

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Tc number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	42	22	AAH76286
2	42	100.0	2596	22	AAH76294
3	42	100.0	2625	22	AAH76283
4	30.4	72.4	42	22	AAH76284
5	30.4	72.4	2334	22	AAH76281
6	30	71.4	1811	20	AAH36522
7	30	71.4	1811	22	AAH76296
8	30	71.4	1811	22	AAH95714
9	30	71.4	1844	22	AAH55663

c	10	30	71.4	2295	22	AAH44653	Human full-length
	11	24.4	58.1	42	22	AAH76285	PAX8e9-PPARGammae1
	12	24.4	58.1	2523	22	AAH76282	Human PAX8e9-PPARG
	13	24.2	57.6	1518	22	AAH76282	Human peroxisome p
	14	24.2	57.6	1608	22	AAH76288	Human PPARGamma CD
	15	24.2	57.6	1608	22	AAH76288	Human PPARGamma co
	16	24.2	57.6	1647	17	AAH35334	Peroxisome prolif
	17	24	57.1	695	20	AAH19066	Human PPARGamma i
	18	23.6	56.2	33	21	AAH257563	Human PPARGamma p
	19	23.6	56.2	33	22	AAH26111	Human PPARGamma g
	20	23.6	56.2	33	22	AAH41855	Human PPARGamma p
	21	23.6	56.2	33	22	AAH42011	Body weight gain i
c	22	23.6	56.2	10524	23	AAH59593	Propionibacterium
	23	23	54.8	26	20	AAH00271	Human PPARGamma p
	24	23	54.8	2757	22	AAH16672	Human CDNA sequenc
	25	22.4	53.3	348	22	AAH76306	Human PPARGamma-PA
	26	22.4	53.3	537	22	AAH76305	Human PPARGamma-PA
c	27	22.4	53.3	683	24	ABL83283	Human ovarian canc
	28	22.4	53.3	1372	22	AAH76287	Human PAX8 CDNA se
	29	22.4	53.3	2711	22	AAH76295	Human PAX8 CDNA se
	30	22	52.4	346	24	ABL37401	Human colon tumour
	31	22	52.4	611	21	AAH98099	Human colon cancer
	32	22	52.4	1936	17	AAH35333	Peroxisome prolif
c	33	21.8	51.9	43599	24	ABK84242	Human CDNA differe
	34	21.6	51.4	634	24	ABN63938	Human cancer relat
	35	21.6	51.4	641	24	ABN63434	Human cancer relat
	36	21.6	51.4	6565	23	ABL01922	Drosophila melanog
	37	21.6	51.4	10425	23	ABL13480	Drosophila melanog
	38	21.6	51.4	10426	23	ABL19870	Bovine PPARG coding
	39	21.4	51.0	239	22	AAH1676	Rat PPARGamma A/B
	40	21.4	51.0	277	19	AAH21208	Mouse PPARGamma A
	41	21.4	51.0	373	19	AAH21207	Human type IV coll
c	42	21.4	51.0	684	21	AAH257162	DNA encoding human
c	43	21.4	51.0	684	24	ABK15362	Type IV collagen N
c	44	21.4	51.0	900	20	AAH20090	Human alpha2(IV)NC
c	45	21.4	51.0	900	21	AAH90992	

ALIGNMENTS

RESULT 1
AAH76286
ID AAH76286 standard; DNA; 42 BP.
XX AAH76286;
AC AAH76286;
XX 29-OCT-2001 (first entry)
DT
XX
DE : PAX8e9-PPARGammae1 fusion junction DNA sequence.
XX
KW PAX8-PPARGammae1; oncogene; cytostatic; PAX8; PPARGammae1; cancer;
KW follicular carcinoma; PAX8e9-PPARGammae1; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 3..41
FT /*tag= a
XX
XX WO200152789-A2.
XX
XX 26-JUL-2001.
XX
XX 18-JAN-2001; 2001WO-US01664.
XX
XX 20-JAN-2000; 2000US-0177109.
PR
PR 14-AUG-2000; 2000US-0225079.
XX
XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX
XX Kroll TG, Fletcher JA;
XX

DR WPI: 2001-514487/56.
XX P-PSDB; AAB85798.
XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
PT treating certain tumors or cancers, e.g. follicular carcinoma -
XX
XX
XX Claim 21; Page 119; 145pp; English.
XX
CC The invention relates to an oncogene designated PAX8-PPARGamma1 that
CC contains a PAX8 coding region fused to PPARGamma1 coding region. The
CC PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant
CC methodology. A PPARGamma1 ligand or agent is useful for treating a
CC subject having a disorder characterized by the presence of a PAX8-
CC PPARGamma1, where the disorder is cancer, e.g. follicular carcinoma.
CC The PAX8-PPARGamma1 molecules are also useful for providing nucleotide
CC and amino acid sequences useful for detecting the above disease. The
CC present sequence represents the nucleotide sequence including and
CC surrounding the fusion junction in the PAX8e9-PPARGamma1 DNA.
XX
SQ Sequence 42 BP; 12 A; 10 C; 13 G; 7 T; 0 other;

Query Match 100.0%; Score 42; DB 22; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATCGCAGCATGGTGCAGAAATGACCATGTTGACACAG 42
|||||
DB 1 CCATCGCAGCATGGTGCAGAAATGACCATGTTGACACAG 42

RESULT 2
AAH76294
ID AAH76294 standard; cDNA; 2596 BP.
AC AAH76294;
XX
XX 29-OCT-2001 (first entry)
XX
XX Human PAX8e9(-exon 8)-PPARGamma1 cDNA sequence.
XX
XX PAX8-PPARGamma1; oncogene; cytostatic; PAX8; PPARGamma1; cancer;
KW follicular carcinoma; PAX8e9(-exon 8)-PPARGamma1; human; SS.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 161..2596
FT /tag= a
XX
XX WO200152789-A2.
XX
XX 26-JUL-2001.
XX
XX 18-JAN-2001; 2001WO-US01664.
XX
XX 20-JAN-2000; 2000US-0177109.
XX 14-AUG-2000; 2000US-0225079.
XX
XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX
XX Kroll TG, Fletcher JA;
XX
XX WPI: 2001-514487/56.
XX P-PSDB; AAB85801.
XX
XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
PT treating certain tumors or cancers, e.g. follicular carcinoma -
XX
XX Claim 1; Page 127-131; 145pp; English.
XX
XX The invention relates to an oncogene designated PAX8-PPARGamma1 that
CC contains a PAX8 coding region fused to PPARGamma1 coding region. The
CC PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant

CC methodology. A PPARGamma1 ligand or agent is useful for treating a
CC subject having a disorder characterized by the presence of a PAX8-
CC PPARGamma1, where the disorder is cancer, e.g. follicular carcinoma.
CC The PAX8-PPARGamma1 molecules are also useful for providing nucleotide
CC and amino acid sequences useful for detecting the above disease. The
CC present sequence represents a human PAX8e9(-exon 8)-PPARGamma1
XX polypeptide encoding cDNA.
XX
SQ Sequence 2596 BP; 660 A; 745 C; 672 G; 519 T; 0 other;

Query Match 100.0%; Score 42; DB 22; Length 2596;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATCGCAGCATGGTGCAGAAATGACCATGTTGACACAG 42
|||||
DB 1140 CCATCGCAGCATGGTGCAGAAATGACCATGTTGACACAG 1181

RI JULT 3
AAH76283
ID AAH76283 standard; cDNA; 2625 BP.
XX
XX AC AAH76283;
XX
XX 29-OCT-2001 (first entry)
XX
XX Human PAX8e9-PPARGamma1 cDNA sequence.
XX
XX PAX8-PPARGamma1; oncogene; cytostatic; PAX8; PPARGamma1; cancer;
KW follicular carcinoma; PAX8e9-PPARGamma1; human; SS.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 1..2625
FT /tag= a
XX
XX WO200152789-A2.
XX
XX 26-JUL-2001.
XX
XX 18-JAN-2001; 2001WO-US01664.
XX
XX 20-JAN-2000; 2000US-0177109.
XX 14-AUG-2000; 2000US-0225079.
XX
XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX
XX Kroll TG, Fletcher JA;
XX
XX WPI: 2001-514487/56.
XX P-PSDB; AAB85795.
XX
XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
PT treating certain tumors or cancers, e.g. follicular carcinoma -
XX
XX Claim 1; Page 112-116; 145pp; English.
XX
XX The invention relates to an oncogene designated PAX8-PPARGamma1 that
CC contains a PAX8 coding region fused to PPARGamma1 coding region. The
CC PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant
CC methodology. A PPARGamma1 ligand or agent is useful for treating a
CC subject having a disorder characterized by the presence of a PAX8-
CC PPARGamma1, where the disorder is cancer, e.g. follicular carcinoma.
CC The PAX8-PPARGamma1 molecules are also useful for providing nucleotide
CC and amino acid sequences useful for detecting the above disease. The
CC present sequence represents a human PAX8e9-PPARGamma1 polypeptide
CC encoding cDNA.
XX
SQ Sequence 2625 BP; 662 A; 761 C; 646 G; 556 T; 0 other;

Query Match 100.0%; Score 42; DB 22; Length 2625;

```

Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATCGCGCATGTCGAGAAATGACCATGGTTGACACAG 42
Db 1169 CCATCGCGCATGTCGAGAAATGACCATGGTTGACACAG 1210

RESULT 4
AAH76284
ID AAH76284 standard; DNA; 42 BP.
AC AAH76284;
XX
XX
XX 29-OCT-2001 (first entry)
XX
XX
XX PAX8e7-PPARGammael translocation fusion juncture 1 DNA sequence.
XX
XX PAX8-PPARGammael; oncogene; cytostatic; PAX8; PPARGammael; cancer;
KW follicular carcinoma; PAX8e7-PPARGammael; human; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 3..41
FT /*tag= a
FT
XX
XX WO200152789-A2.
XX
XX 26-JUL-2001.
XX
XX 18-JAN-2001; 2001WO-US01664.
XX
XX 20-JAN-2000; 2000US-0177109.
XX
XX 14-AUG-2000; 2000US-0225079.
XX
XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
XX Kroll TG, Fletcher JA;
XX
XX WPI; 2001-514487/56.
XX
XX P-PSDB; AAB85796.
XX
XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
PT treating certain tumors or cancers, e.g. follicular carcinoma
PT
XX
XX Claim 21; Page 118; 145pp; English.
XX
XX The invention relates to an oncogene designated PAX8-PPARGammael that
C contains a PAX8 coding region fused to PPARGammael coding region. The
C PAX8-PPARGammael polypeptides can be expressed by standard recombinant
C methodology. A PPARGammael ligand or agent is useful for treating a
C subject having a disorder characterized by the presence of a PAX8-
C PPARGammael, where the disorder is cancer, e.g. follicular carcinoma.
C The PAX8-PPARGammael molecules are also useful for providing nucleotide
C and amino acid sequences useful for detecting the above disease. The
C present sequence represents the nucleotide sequence including and
C surrounding the translocation fusion juncture in the PAX8e7-PPARGammael
C DNA.
XX
XX Sequence 42 BP; 12 A; 11 C; 12 G; 7 T; 0 other;

Query Match 72.4%; Score 30.4; DB 22; Length 42;
Best Local Similarity 96.9%; Pred. No. 0.0065;
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 CATGTGGCAGAAATGACCATGGTTGACACAG 42
Db 11 CGTGTGGCAGAAATGACCATGGTTGACACAG 42

RESULT 5
AAH76281
ID AAH76281 standard; cDNA; 2334 BP.
AC AAH76281;
XX
XX
XX 29-OCT-2001 (first entry)
XX
XX
XX Human PAX8e7-PPARGammael cDNA sequence.
XX
XX PAX8-PPARGammael; oncogene; cytostatic; PAX8; PPARGammael; cancer;
KW follicular carcinoma; PAX8e7-PPARGammael; human; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 1..2334
FT /*tag= a
FT
XX
XX WO200152789-A2.
XX
XX 26-JUL-2001.
XX
XX 18-JAN-2001; 2001WO-US01664.
XX
XX 20-JAN-2000; 2000US-0177109.
XX
XX 14-AUG-2000; 2000US-0225079.
XX
XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
XX Kroll TG, Fletcher JA;
XX
XX WPI; 2001-514487/56.
XX
XX P-PSDB; AAB85793.
XX
XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
PT treating certain tumors or cancers, e.g. follicular carcinoma
PT
XX
XX Claim 1; Page 100-104; 145pp; English.
XX
XX The invention relates to an oncogene designated PAX8-PPARGammael that
C contains a PAX8 coding region fused to PPARGammael coding region. The
C PAX8-PPARGammael polypeptides can be expressed by standard recombinant
C methodology. A PPARGammael ligand or agent is useful for treating a
C subject having a disorder characterized by the presence of a PAX8-
C PPARGammael, where the disorder is cancer, e.g. follicular carcinoma.
C The PAX8-PPARGammael molecules are also useful for providing nucleotide
C and amino acid sequences useful for detecting the above disease. The
C present sequence represents a human PAX8e7-PPARGammael polypeptide
C encoding cDNA.
XX
XX Sequence 2334 BP; 614 A; 650 C; 573 G; 497 T; 0 other;

Query Match 72.4%; Score 30.4; DB 22; Length 2334;
Best Local Similarity 96.9%; Pred. No. 0.017; 1; Indels 0; Gaps 0;
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 CATGTGGCAGAAATGACCATGGTTGACACAG 42
Db 888 CGTGTGGCAGAAATGACCATGGTTGACACAG 919

RESULT 6
AAH76282
ID AAH76282 standard; cDNA; 1811 BP.
AC AAH76282;
XX
XX
XX 07-JUL-1999 (first entry)
XX
XX
XX Human PPARGammael coding sequence.
XX
XX Nuclear receptor agonist; antagonist; identification; PPARG;
KW peroxisome proliferator activated receptor; ss.
XX
XX

```

OS Homo sapiens.
 PN WO9918124-A1.
 XX
 PD 15-APR-1999.
 XX
 PF 06-OCT-1998; 98WO-US21049.
 XX
 PR 07-OCT-1997; 97US-0061385.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 PI Cummings RT, Hermes JD, Moller DE, Zhou G;
 XX WPI; 1999-263998/22.
 DR P-PSDB; AAY05471.
 XX
 PT Identifying nuclear receptor agonists and antagonists
 XX
 PS Disclosure; Fig 9b; 60pp; English.
 CC This sequence encodes the human peroxisome proliferator activated
 CC receptor-gamma (PPAR-gamma).
 CC The invention relates to a method for identifying nuclear receptor
 CC agonists and antagonists comprises measuring fluorescent resonance energy
 CC transfer between fluorescent-labelled nuclear receptors and
 CC co-activators. The method can be used for identifying agonists and
 CC antagonist of nuclear receptors.
 XX
 SQ Sequence 1811 BP; 510 A; 433 C; 422 G; 446 T; 0 other;
 Query Match 71.4%; Score 30; DB 20; Length 1811;
 Best Local Similarity 86.8%; Pred. No. 0.023;
 Matches 33; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 5 CGCAGGCATGGTGGCAGAAATGACCATGGTTGACACAG 42
 |||||
 DB 154 CGCCGCCGTGGCGCAGAAATGACCATGGTTGACACAG 191
 |||||
 RESULT 7
 AAH76296
 ID AAH76296 standard; cDNA; 1811 BP.
 XX
 AC AAH76296;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Human PPARGamma cDNA sequence.
 XX
 KW PAX8-PPARGamma; oncogene; cytostatic; PAX8; PPARGamma; cancer;
 KW follicular carcinoma; PPARGamma; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS /tag= a
 FT
 XX WO200152789-A2.
 PN
 PD 26-JUL-2001.
 XX
 PF 18-JAN-2001; 2001WO-US01664.
 XX
 PR 20-JAN-2000; 2000US-0177109.
 PR 14-AUG-2000; 2000US-0225079.
 XX
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
 XX
 PI Kroll TG, Fletcher JA;
 XX WPI; 2001-514487/56.
 DR

DR P-PSDB; AAB85802.
 XX
 PT New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
 PT treating certain tumors or cancers, e.g. follicular carcinoma
 XX
 PS Disclosure; Page 137-139; 145pp; English.
 XX
 CC The invention relates to an oncogene designated PAX8-PPARGamma1 that
 CC contains a PAX8 coding region fused to PPARGamma1 coding region. The
 CC PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant
 CC methodology. A PPARGamma ligand or agent is useful for treating a
 CC subject having a disorder characterized by the presence of a PAX8-
 CC PPARGamma1, where the disorder is cancer, e.g. follicular carcinoma.
 CC The PAX8-PPARGamma1 molecules are also useful for providing nucleotide
 CC and amino acid sequences useful for detecting the above disease. The
 CC present sequence represents a human PPARGamma polypeptide encoding cDNA.
 XX
 SQ Sequence 1811 BP; 510 A; 433 C; 422 G; 446 T; 0 other;
 Query Match 71.4%; Score 30; DB 22; Length 1811;
 Best Local Similarity 86.8%; Pred. No. 0.023;
 Matches 33; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 5 CGCAGGCATGGTGGCAGAAATGACCATGGTTGACACAG 42
 |||||
 DB 154 CGCCGCCGTGGCGCAGAAATGACCATGGTTGACACAG 191
 |||||
 RESULT 8
 ABN95714
 ID ABN95714 standard; DNA; 1811 BP.
 XX
 AC ABN95714;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Gene #2212 used to diagnose liver cancer.
 XX
 KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 KW metastatic liver tumour; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX
 OS Homo sapiens.
 XX
 PN WO200229103-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 02-OCT-2001; 2001WO-US30589.
 XX
 PR 02-OCT-2000; 2000US-237054P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 XX WPI; 2002-426119/45.
 XX
 PT Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample -
 XX
 PS Claim 1; SEQ ID NO 2212; 298pp; English.
 XX
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic

CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1811 BP; 510 A; 433 C; 422 G; 446 T; 0 other;

Query Match 71.4%; Score 30; DB 24; Length 1811;

Best Local Similarity 86.8%; Pred. No. 0.023; Mismatches 0; Gaps 0;

Matches 33; Conservative 0; Indels 5; Indels 0; Gaps 0;

QY 5 CGCAGCGCATGGTGGCAGAAATGACCATGGTTGACACAG 42
 |||||
 Db 154 CGCCGCGTGGCCGCGAGAAATGACCATGGTTGACACAG 191

RESULT 9

AJ 563
 IL AAF55663 standard; DNA; 1844 BP.

AC AAF55663;

XX

DT 29-MAY-2001 (first entry)

XX

DE Human peroxisome proliferator-activated receptor gamma DNA.

XX

KW Human; peroxisome proliferator-activated receptor gamma; PPAR-gamma;

KW orphan receptor; cancer; ss.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT CDS 179..1603

FT /*tag= a

FT /product= "peroxisome proliferator-activated receptor

FT gamma"

XX US6200802-B1.

XX 13-MAR-2001.

XX

PF 09-OCT-1993; 93US-0134557.

XX

PR 08-OCT-1993; 93US-0134557.

XX

XX (ARCH-) ARCH DEV CORP.

F

P1 Greene ME, Blumberg B;

XX

DR WPI; 2001-234517/24.

DR P-PSDB; AAB67656.

XX

PT Novel isolated and purified polynucleotide encoding human peroxisome

PT proliferator-activated receptor gamma polypeptide useful in

PT diagnostics, drug design and therapeutics

XX

PS Claim 3; Column 43-48; 29pp; English.

XX

CC The present sequence encodes a human peroxisome proliferator-activated

CC receptor gamma (PPAR-gamma) polypeptide. PPAR are orphan receptors, and

CC may play a role in proliferative and differentiation aspects of cancer.

CC The PPAR-gamma polynucleotide is useful for detecting a mRNA transcript

CC that encodes PPAR polypeptide. It is also useful for detecting hybrid

CC formation. The PPAR-gamma polypeptide is useful in diagnostics, drug

CC design and therapeutics.

XX

XX Sequence 1844 BP; 543 A; 433 C; 421 G; 447 T; 0 other;

SQ

Query Match

Best Local Similarity 71.4%; Score 30; DB 22; Length 1844;

Mismatches 5; Indels 0; Gaps 0;

Matches 33; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 CGCAGCGCATGGTGGCAGAAATGACCATGGTTGACACAG 42

|||||

Db 154 CGCCGCGTGGCCGCGAGAAATGACCATGGTTGACACAG 191

RESULT 10

AAS44653/c

ID AAS44653 standard; DNA; 2295 BP.

XX

AC AAS44653;

XX

DT 18-DEC-2001 (first entry)

XX

XX Human full-length polynucleotide sequence #78.

XX

KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;

KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;

KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;

KW nervous system disorder; inflammatory disorder; cell differentiation; ds;

KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;

KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;

KW cytostatic; antirheumatic; antiarthritic; vulnery; antinflammatory;

KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;

KW neuroprotective; osteopathic; antidiabetic; antiallergic;

KW immunostimulant; analgesic; gene therapy.

XX

OS Homo sapiens.

XX

PN WO200164834-A2.

XX

XX 07-SEP-2001.

XX

XX 26-FEB-2001; 2001WO-US04926.

XX

XX 28-FEB-2000; 2000US-0515126.

XX

XX 18-MAY-2000; 2000US-0577409.

XX

XX 17-JUN-2000; 2000US-0597707.

XX

XX 14-JUL-2000; 2000US-0616807.

XX

XX 19-SEP-2000; 2000US-0664641.

XX

XX (HYSE-) HYSEQ INC.

XX

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao CA, Ren F;

XX

XX Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;

XX

XX Drmanac R;

XX

XX WPI; 2001-589962/66.

XX

XX P-PSDB; AAU27753.

XX

XX Novel polypeptides and nucleic acids obtained from cDNA libraries

XX prepared from various human tissues, for diagnosis, treatment of

XX cancer, neurological, inflammatory disorders and for use in arrays for

XX detection

XX

XX Claim 1; SEQ ID No 78; 153pp; English.

XX

XX Sequences AAS44576-AAS44919 represent full-length polynucleotides and

XX contig polynucleotides encoding polypeptides of the invention. The DNA

XX and protein sequences are useful for the treatment, diagnosis and

XX prevention of various types of disorder in a mammalian subject such as a

XX human, dog, monkey, mouse, hamster or rat. The disorders include cancers

XX such as leukemia, lymphoma and neuroblastoma, autoimmune disorders such

XX as multiple sclerosis, connective tissue disease, rheumatoid arthritis,

XX diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system

XX disorders such as Parkinson's disease, Alzheimer's disease, Huntington's

XX chorea, ankyrotrophic lateral sclerosis, spinal muscular atrophy and

XX wernicke disease, inflammatory disorders such as nephritis, Crohn's

XX disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory

XX bowel disease. The sequences exhibit activity relating to angiogenesis,

XX cell proliferation, cell differentiation, stem cell growth factor,

XX activin or inhibin. Therefore, they can be used to manipulate stem cells

Query Match 58.1%; Score 24.4; DB 22; Length 42;

AAD21022
 ID AAD21022 standard; cDNA; 1518 BP.


```

XX (ISIS-) ISIS PHARM INC.
PA
XX PI
XX McKay R, Baker BZ, Borchers AH;
XX WPI; 2001-070112/08.
DR P-PSDB; AAB59839.
XX
PT Novel antisense compounds capable of modulating expression of
PT peroxisome proliferator-activated receptor gamma useful for diagnosis,
PT prophylaxis and treatment of diseases associated with expression of the
PT receptor
XX
PS Example 13; Column 45-48; 40pp; English.
XX
CC Peroxisome proliferator-activated receptors (PPARs) are members of the
CC nuclear hormone receptor subfamily of transcription factors. The present
CC invention relates to antisense oligonucleotides (see AAF23652-F23731),
CC targeted to a nucleic acid molecule encoding human PPARgamma, which
CC specifically hybridizes with and inhibits the expression of human
CC PPARgamma. The present sequence is the coding sequence for human
CC PPARgamma. The PPARgamma antisense oligonucleotides of the present
CC invention can be used in the diagnosis and treatment of diseases
CC associated with the expression of PPARgamma, e.g. to prevent or delay
CC infection, inflammation or tumour formation.
XX
SQ Sequence 1608 BP; 462 A; 380 C; 361 G; 405 T; 0 other;
Query Match 57.6%; Score 24.2; DB 22; Length 1608;
Best Local Similarity 78.4%; Pred. No. 5.1;
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 6 GCAGGCGATGTCGACAAATGACCATGGTTGACACAG 42
Db 157 GCAACATATATCAAGAAATGACCATGGTTGACACAG 193
Search completed: January 21, 2003, 23:09:56
Job time : 129.667 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 23:03:25 ; Search time 26.3333 Seconds
(without alignments)
489.130 Million cell updates/sec

Title: US-09-765-111A-11
Perfect score: 42
Sequence: 1 ccacgcagcatgtggca.....aatgaccatggtgacacag 42

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Tc number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	71.4	1844	US-08-134-557D-1	Sequence 1, Appli
2	24.2	57.6	1518	US-09-128-142-3	Sequence 3, Appli
3	24.2	57.6	1608	US-09-484-345-3	Sequence 3, Appli
4	24.2	57.6	1679	US-09-514-247A-5	Sequence 5, Appli
5	21.4	51.0	277	US-08-917-653-4	Sequence 4, Appli
6	21.4	51.0	373	US-08-917-653-3	Sequence 3, Appli
7	21.4	51.0	900	US-09-589-927-3	Sequence 3, Appli
8	21.4	51.0	900	US-09-277-665-3	Sequence 3, Appli
9	21.4	51.0	4359	US-09-484-970B-4	Sequence 4, Appli
10	21.2	50.5	34303	US-08-735-609-4	Sequence 4, Appli
11	21.2	50.5	34303	US-08-735-609-4	Sequence 4, Appli
12	21.2	50.5	34303	US-09-315-372-4	Sequence 4, Appli
13	21.2	50.5	34303	US-09-244-752-4	Sequence 4, Appli
14	21.2	50.5	34303	US-09-245-497-4	Sequence 4, Appli
15	21.2	50.5	34303	US-09-562-919-4	Sequence 4, Appli
16	21.2	50.5	34382	US-08-374-483-6	Sequence 6, Appli
17	21.2	50.5	35408	US-08-973-334-3	Sequence 3, Appli
18	21.2	50.5	35408	US-09-563-869A-3	Sequence 3, Appli
19	21.2	50.5	35408	US-08-549-489-3	Sequence 3, Appli
20	21.2	50.5	35935	US-08-735-609-1	Sequence 1, Appli
21	21.2	50.5	35935	US-08-735-609-1	Sequence 1, Appli
22	21.2	50.5	35935	US-08-379-452-43	Sequence 43, Appli
23	21.2	50.5	35935	US-09-315-372-1	Sequence 1, Appli
24	21.2	50.5	35935	US-09-244-752-1	Sequence 1, Appli
25	21.2	50.5	35935	US-09-245-497-1	Sequence 1, Appli
26	21.2	50.5	35935	US-09-409-670-43	Sequence 43, Appli
27	21.2	50.5	35935	US-09-562-919-1	Sequence 1, Appli

28	20.8	49.5	1796	4	US-09-255-392-1	Sequence 1, Appli
29	20.8	49.5	2005	2	US-08-484-200-1	Sequence 1, Appli
30	20.8	49.5	2005	2	US-08-477-493-1	Sequence 1, Appli
31	20.8	49.5	2005	3	US-08-465-375-1	Sequence 1, Appli
32	20.8	49.5	2005	4	US-09-788-070-1	Sequence 1, Appli
C 33	20.6	49.0	1134	4	US-09-206-059-29	Sequence 29, Appli
C 34	20.6	49.0	2296	1	US-07-750-080A-18	Sequence 18, Appli
C 35	20.6	49.0	2296	3	US-08-651-472-18	Sequence 18, Appli
C 36	20.6	49.0	2296	4	US-08-358-928-18	Sequence 18, Appli
C 37	20.6	49.0	2497	1	US-08-643-219-12	Sequence 12, Appli
C 38	20.6	49.0	2497	2	US-09-131-995-12	Sequence 12, Appli
C 39	20.6	49.0	2497	2	US-08-832-087B-12	Sequence 12, Appli
C 40	20.6	49.0	2497	3	US-08-851-350-12	Sequence 12, Appli
C 41	20.6	49.0	2497	4	US-09-132-154-12	Sequence 12, Appli
C 42	20.6	49.0	2679	6	5200340-7	Patent No. 5200340
C 43	20.6	49.0	2753	1	US-07-854-603-1	Sequence 1, Appli
44	20.2	48.1	1977	4	US-09-595-424-5	Sequence 5, Appli
45	20.2	48.1	3258	4	US-09-595-424-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-08-134-557D-1
; Sequence 1, Application US/08134557D
; Patent No. 6200802
; GENERAL INFORMATION:
; APPLICANT: Greene, Marianne E.
; APPLICANT: Blumberg, Bruce
; TITLE OF INVENTION: Human Peroxisome Proliferator Activated
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rocky, Milnamow & Katz, Ltd.
; STREET: 2 Prudential Plaza, Suite 4700 180 N. Stetson
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/134,557D
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Katz, Martin L.
; REGISTRATION NUMBER: 25,011
; REFERENCE/DOCKET NUMBER: ARCH:098
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5400
; TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1844 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 179..1606
; US-08-134-557D-1

Query Match 71.4%; Score 30; DB 4; Length 1844;
Best Local Similarity 86.8%; Pred. No. 0.008; Indels 0; Gaps 0;
Matches 33; Conservative 0; Mismatches 5;
Qy 5 CGCAGCATGTGGCAGAAATGACCATGTTGACACAG 42

Db 154 CGCGCGTGGCGGAGAAATGACCATGGTTGACACAG 191
|||||
RESULT 2
US-09-128-142-3
; Sequence 3, Application US/09128142
; Patent No. 6294559
; GENERAL INFORMATION:
; APPLICANT: Smith, Roy G.
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENTS ASSOCIATED WITH
; PEROXISOME PROLIFERATOR ACTIVATED RECEPTORS GAMMAL AND GAMM
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jack L. Tribble
; STREET: 126 E. Lincoln Ave., P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Power Macintosh 7500/100
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128,142
; FILING DATE: 03-Aug-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,007
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Tribble, Jack L.
; REGISTRATION NUMBER: 32,633
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-5321
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1518 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-128-142-3
Query Match 57.6%; Score 24.2; DB 4; Length 1518;
Best Local Similarity 78.4%; Pred. No. 1.3;
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 6 GCAGGCATGGTGGCAGAAATGACCATGGTTGACACAG 42
Db 67 GCAACATATCACAAGAAATGACCATGGTTGACACAG 103
|||||
RESULT 3
US-09-484-345-3
; Sequence 3, Application US/09484345
; Patent No. 6159734
; GENERAL INFORMATION:
; APPLICANT: Robert McKay
; APPLICANT: Alexander H. Borchers
; APPLICANT: Brenda F. Baker
; TITLE OF INVENTION: ANTISENSE MODULATION OF PEROXISOME PROLIFERATOR-ACTIVATED RECEPT
; FILE REFERENCE: RTS-0104
; CURRENT APPLICATION NUMBER: US/09/484,345
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 3
; LENGTH: 1608

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (91)...(1608)
US-09-484-345-3
Query Match 57.6%; Score 24.2; DB 3; Length 1608;
Best Local Similarity 78.4%; Pred. No. 1.3;
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 6 GCAGGCATGGTGGCAGAAATGACCATGGTTGACACAG 42
Db 157 GCAACATATCACAAGAAATGACCATGGTTGACACAG 193
|||||
RESULT 4
US-09-514-247A-5
; Sequence 5, Application US/09514247A
; Patent No. 6365361
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO. LTD.
; APPLICANT: TANIGUCHI, Tomoyasu
; APPLICANT: MIZUKAMI, Junko
; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PP
; FILE REFERENCE: TANIGUCHI=6
; CURRENT APPLICATION NUMBER: US/09/514,247A
; CURRENT FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/JP98/03734
; PRIOR FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: JP231084/1997
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)...(1679)
US-09-514-247A-5
Query Match 57.6%; Score 24.2; DB 4; Length 1679;
Best Local Similarity 78.4%; Pred. No. 1.4;
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 6 GCAGGCATGGTGGCAGAAATGACCATGGTTGACACAG 42
Db 225 GCAACATATCACAAGAAATGACCATGGTTGACACAG 261
|||||
RESULT 5
US-08-917-653-4
; Sequence 4, Application US/08917653
; Patent No. 6004751
; GENERAL INFORMATION:
; APPLICANT: Rosenfield, Robert L.
; TITLE OF INVENTION: IDENTIFICATION OF ACTIVATORS AND
; TITLE OF INVENTION: INHIBITORS OF SEBUM FORMATION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

APPLICANT: Walker, Michael G.
TITLE OF INVENTION: BONE REMODELING GENES
FILE REFERENCE: PB-0014 US
CURRENT APPLICATION NUMBER: US/09/484,970B
CURRENT FILING DATE: 2000-01-18
NUMBER OF SEQ ID NOS: 172
SEQ ID NO 4
SOFTWARE: PERL Program
TYPE: DNA
LENGTH: 4359
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. 6426186 331365 (1553795CB1)
US-09-484-970B-4

Query Match 51.0%; Score 21.4; DB 4; Length 4359;
Best Local Similarity 80.6%; Pred. No. 19;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 6 GCAGGCGATGGTGGCAGAAATGACCATGGTTG 36
DB 3054 GTGCGCATAGTAGCAGACATCCAGGGTTG 3024

RESULT 10
US-08-735-609-4
Sequence 4, Application US/08735609
Patent No. 5955360
GENERAL INFORMATION:
APPLICANT: Chamberlain, Jeffrey S.
APPLICANT: Amalfitano, Andrea
APPLICANT: Hauser, Michael A.
APPLICANT: Kumar-Singh, Rajendra
APPLICANT: Hartigan-O'Connor, Dennis J.
TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/735,609
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: UM-02484
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 34303 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
US-08-735-609-4

Query Match 50.5%; Score 21.2; DB 2; Length 34303;
Best Local Similarity 69.0%; Pred. No. 34;
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CCATCGCAGGCGATGGTGGCAGAAATGACCATGGTTGACACAG 42
DB 18736 CCACAGAGGCGATGGAGACACAAACGTCGCCGGTTGCCTCAG 18777

RESULT 11
US-08-735-609-4
Sequence 4, Application US/08735609
Patent No. 5994132
GENERAL INFORMATION:
APPLICANT: Chamberlain, Jeffrey S.
APPLICANT: Amalfitano, Andrea
APPLICANT: Hauser, Michael A.
APPLICANT: Kumar-Singh, Rajendra
APPLICANT: Hartigan-O'Connor, Dennis J.
TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/735,609
FILING DATE: 23-Oct-1996
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: UM-02484
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 34303 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
US-08-735-609-4

Query Match 50.5%; Score 21.2; DB 2; Length 34303;
Best Local Similarity 69.0%; Pred. No. 34;
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 CCATCGCAGGCGATGGTGGCAGAAATGACCATGGTTGACACAG 42
DB 18736 CCACAGAGGCGATGGAGACACAAACGTCGCCGGTTGCCTCAG 18777

RESULT 12
US-09-315-372-4
Sequence 4, Application US/09315372
Patent No. 6057158
GENERAL INFORMATION:
APPLICANT: Chamberlain, Jeffrey S.
APPLICANT: Amalfitano, Andrea
APPLICANT: Hauser, Michael A.
APPLICANT: Kumar-Singh, Rajendra
APPLICANT: Hartigan-O'Connor, Dennis J.
TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
NUMBER OF SEQUENCES: 15


```
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
; US-09-245-497-4

Query Match          50.5%; Score 21.2; DB 3; Length 34303;
Best Local Similarity 69.0%; Pred. NO. 34;
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 CCATCCAGGCGATGGTGCAGAAATGACCATGCTTGACACAG 42
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 18736 CCACAGGGCGATGGAGACACAAACGTCGCCGTTGCCTCAG 18777

RESULT 15
US-09-562-919-4
; Sequence 4, Application US/09562919
; Patent No. 6451596
; GENERAL INFORMATION:
; APPLICANT: Chamberlain, Jeffrey S.
; Amalfitano, Andrea
; Hauser, Michael A.
; Kumar-Singh, Rajendra
; Hartigan-O'Connor, Dennis J.
; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/562,919
; FILING DATE: 02-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/735,609
; FILING DATE: 23-Oct-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: UM-02484
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34303 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-562-919-4

Query Match          50.5%; Score 21.2; DB 4; Length 34303;
Best Local Similarity 69.0%; Pred. NO. 34;
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 CCATCCAGGCGATGGTGCAGAAATGACCATGCTTGACACAG 42
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 18736 CCACAGGGCGATGGAGACACAAACGTCGCCGTTGCCTCAG 18777

Search completed: January 22, 2003, 00:43:47
Job time : 59.3333 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 23:59:20 ; Search time 27 Seconds
(without alignments)
693.573 Million cell updates/sec

Title: US-09-765-111A-11
Perfect score: 42
Sequence: 1 ccacgcaggcatggtgcga.....aatgaccatggtgacacag 42

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 393868 seqs, 222934149 residues
Tc number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	42	10	US-09-765-111A-11
2	42	100.0	2596	10	US-09-765-111A-22
3	42	100.0	2625	10	US-09-765-111A-5
4	30.4	72.4	42	10	US-09-765-111A-7
5	30.4	72.4	2334	10	US-09-765-111A-1
6	30	71.4	1811	10	US-09-765-111A-26
7	30	71.4	1811	10	US-09-880-107-2212
8	30	71.4	2260	10	US-09-816-828-8
9	30	71.4	2329	10	US-09-816-828-9
10	24.4	58.1	42	10	US-09-765-111A-9
11	24.4	58.1	2523	10	US-09-765-111A-3
12	24.2	57.6	1608	10	US-09-765-111A-15
13	24.2	57.6	1679	12	US-10-109-886-5
14	22.4	53.3	348	10	US-09-765-111A-37
15	22.4	53.3	537	10	US-09-765-111A-36
16	22.4	53.3	683	10	US-09-867-701-6261
17	22.4	53.3	1372	10	US-09-765-111A-13
18	22.4	53.3	2711	10	US-09-765-111A-24
19	22	52.4	343	9	US-10-046-935-990

Sequence 990, Appl
Sequence 109, Appl
Sequence 262, Appl
Sequence 3682, Appl
Sequence 75, Appl
Sequence 23, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 1, Appl
Sequence 5, Appl
Sequence 1, Appl
Sequence 7349, Appl
Sequence 7320, Appl
Sequence 1, Appl
Sequence 20992, A
Sequence 4239, Appl
Sequence 14476, A
Sequence 612, Appl
Sequence 1, Appl
Sequence 611, Appl
Sequence 617, Appl
Sequence 1605, Appl
Sequence 4, Appl
Sequence 1675, Appl
Sequence 6, Appl

20 22 52.4 343 9 US-09-878-178-990
21 22 52.4 611 10 US-09-925-299-109
22 21.6 51.4 3453 9 US-10-108-605-262
23 21.4 51.0 2212 10 US-09-880-107-3682
24 21.4 51.0 3233 10 US-09-925-302-75
25 21.2 50.5 32480 9 US-09-847-101B-23
26 21.2 50.5 35871 10 US-09-956-335-2
27 21.2 50.5 35935 10 US-09-725-720-43
28 21.2 50.5 35935 10 US-09-782-378A-4
29 21.2 50.5 35935 10 US-09-782-378A-5
30 21.2 50.5 35978 10 US-09-956-335-1
31 21 50.0 431 10 US-09-867-701-7349
32 21 50.0 449 10 US-09-867-701-7320
33 20.8 49.5 2005 12 US-10-143-373-1
34 20.6 49.0 160 10 US-09-864-761-20992
35 20.6 49.0 362 10 US-09-864-761-42339
36 20.6 49.0 395 10 US-09-960-352-14476
37 20.6 49.0 1134 10 US-09-873-676-29
38 20.6 49.0 1547 10 US-09-764-870-612
39 20.6 49.0 2497 10 US-09-946-893-1
40 20.6 49.0 32195 10 US-09-764-870-611
41 20.6 49.0 32195 10 US-09-764-870-617
42 20.6 49.0 32195 10 US-09-764-869-1605
43 20.4 48.6 1175 9 US-10-051-307-4
44 20.4 48.6 1459 10 US-09-917-800A-1675
45 20.4 48.6 2945 9 US-09-978-295A-6

ALIGNMENTS

RESULT 1
US-09-765-111A-11
; Sequence 11, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PPARGAMMA NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(41)
US-09-765-111A-11

Query Match 100.0%; Score 42; DB 10; Length 42;
Best Local Similarity 100.0%; Pred. No. 6.9e-08;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATCGAGGATGTCGAGAAATGACCATGGTTGACACAG 42
Db 1 CCATCGAGGATGTCGAGAAATGACCATGGTTGACACAG 42

RESULT 2
US-09-765-111A-22
; Sequence 22, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.

APPLICANT: Kroll, Todd G.
TITLE OF INVENTION: PAX8-PPARGamma NUCLEIC ACID MOLECULES
AND POLYPEPTIDES AND USES THEREOF
FILE REFERENCE: B0801/7196/ERP/MAT
CURRENT APPLICATION NUMBER: US/09/765,111A
CURRENT FILING DATE: 2001-01-18
PRIOR APPLICATION NUMBER: US 60/177,109
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: US 60/225,079
PRIOR FILING DATE: 2000-08-14
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 22
LENGTH: 2596
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (161)...(2596)
US-09-765-111A-22

Query Match 100.0%; Score 42; DB 10; Length 2596;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATCGCAGGCATGGTGCAGAAATGACCATGGTTGACACAG 42
|||||
Db 1140 CCATCGCAGGCATGGTGCAGAAATGACCATGGTTGACACAG 1181

RESULT 3

US-09-765-111A-5
Sequence 5, Application US/09765111A
Patent No. US20020106796A1

GENERAL INFORMATION:
APPLICANT: Fletcher, Jonathan A.
APPLICANT: Kroll, Todd G.
TITLE OF INVENTION: PAX8-PPARGamma NUCLEIC ACID MOLECULES
AND POLYPEPTIDES AND USES THEREOF
FILE REFERENCE: B0801/7196/ERP/MAT
CURRENT APPLICATION NUMBER: US/09/765,111A
CURRENT FILING DATE: 2001-01-18
PRIOR APPLICATION NUMBER: US 60/177,109
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: US 60/225,079
PRIOR FILING DATE: 2000-08-14
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 2625
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(2625)
US-09-765-111A-5

Query Match 100.0%; Score 42; DB 10; Length 2625;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATCGCAGGCATGGTGCAGAAATGACCATGGTTGACACAG 42
|||||
Db 1169 CCATCGCAGGCATGGTGCAGAAATGACCATGGTTGACACAG 1210

RESULT 4

US-09-765-111A-7
Sequence 7, Application US/09765111A
Patent No. US20020106796A1

GENERAL INFORMATION:
APPLICANT: Fletcher, Jonathan A.
APPLICANT: Kroll, Todd G.

TITLE OF INVENTION: PAX8-PPARGamma NUCLEIC ACID MOLECULES
AND POLYPEPTIDES AND USES THEREOF
FILE REFERENCE: B0801/7196/ERP/MAT
CURRENT APPLICATION NUMBER: US/09/765,111A
CURRENT FILING DATE: 2001-01-18
PRIOR APPLICATION NUMBER: US 60/177,109
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: US 60/225,079
PRIOR FILING DATE: 2000-08-14
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 42
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (3)...(41)
US-09-765-111A-7

Query Match 72.4%; Score 30.4; DB 10; Length 42;
Best Local Similarity 96.9%; Pred. No. 0.0026;
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 CATGGTGGCAGAAATGACCATGGTTGACACAG 42
|||||
Db 11 CGTGGTGGCAGAAATGACCATGGTTGACACAG 42

RESULT 5

US-09-765-111A-1
Sequence 1, Application US/09765111A
Patent No. US20020106796A1

GENERAL INFORMATION:
APPLICANT: Fletcher, Jonathan A.
APPLICANT: Kroll, Todd G.
TITLE OF INVENTION: PAX8-PPARGamma NUCLEIC ACID MOLECULES
AND POLYPEPTIDES AND USES THEREOF
FILE REFERENCE: B0801/7196/ERP/MAT
CURRENT APPLICATION NUMBER: US/09/765,111A
CURRENT FILING DATE: 2001-01-18
PRIOR APPLICATION NUMBER: US 60/177,109
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: US 60/225,079
PRIOR FILING DATE: 2000-08-14
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2334
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(2334)
US-09-765-111A-1

Query Match 72.4%; Score 30.4; DB 10; Length 2334;
Best Local Similarity 96.9%; Pred. No. 0.0065;
Matches 31; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 CATGGTGGCAGAAATGACCATGGTTGACACAG 42
|||||
Db 888 CGTGGTGGCAGAAATGACCATGGTTGACACAG 919

RESULT 6

US-09-765-111A-26
Sequence 26, Application US/09765111A
Patent No. US20020106796A1

GENERAL INFORMATION:
APPLICANT: Fletcher, Jonathan A.
APPLICANT: Kroll, Todd G.
TITLE OF INVENTION: PAX8-PPARGamma NUCLEIC ACID MOLECULES

; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 1811
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (173)...(1609)
US-09-765-111a-26

Query Match 71.4%; Score 30; DB 10; Length 1811;
Local Similarity 86.8%; Pred. No. 0.0089;
Matches 33; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 CCGAGCATGTCGCAGAAATGACCATGTTGACACAG 42
Db 154 CCGCCCGTCGCGCAGAAATGACCATGTTGACACAG 191

RESULT 7

US-09-880-107-2212
; Sequence 2212, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2212
; LENGTH: 1811
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L40904
US-09-880-107-2212

Query Match 71.4%; Score 30; DB 10; Length 1811;
Best Local Similarity 86.8%; Pred. No. 0.0089;
Matches 33; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 CCGAGCATGTCGCAGAAATGACCATGTTGACACAG 42
Db 154 CCGCCCGTCGCGCAGAAATGACCATGTTGACACAG 191

RESULT 8

US-09-816-828-8/c
; Sequence 8, Application US/09816828
; Patent No. US20020150898A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod

; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhao, Qing A.
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 791CIP2E
; CURRENT APPLICATION NUMBER: US/09/816,828
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 8
; LENGTH: 2250
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1612)..(2142)
US-09-816-828-8

Query Match 71.4%; Score 30; DB 10; Length 2260;
Best Local Similarity 86.8%; Pred. No. 0.0093;
Matches 33; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 CCGAGCATGTCGCAGAAATGACCATGTTGACACAG 42
Db 1658 CCGCCCGTCGCGCAGAAATGACCATGTTGACACAG 1621

RESULT 9

US-09-816-828-9/c
; Sequence 9, Application US/09816828
; Patent No. US20020150898A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhao, Qing A.
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 791CIP2E
; CURRENT APPLICATION NUMBER: US/09/816,828
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 9
; LENGTH: 2329
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1612)..(2211)
US-09-816-828-9

Query Match 71.4%; Score 30; DB 10; Length 2329;
Best Local Similarity 86.8%; Pred. No. 0.0094;
Matches 33; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 CGCAGGCGTGGCGAGAAATGACCATGGTTGACACAG 42
DB 1658 CGCGCGGTGGCGAGAAATGACCATGGTTGACACAG 1621

RESULT 10

US-09-765-111A-9
; Sequence 9, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PPARGgamma NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(41)
US-09-765-111A-9

Query Match 58.1%; Score 24.4; DB 10; Length 42;
Best Local Similarity 82.4%; Pred. No. 0.6;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 9 GGCATGCTGGCAGAAATGACCATGGTTGACACAG 42
DB 9 GCCTCTCTCAGAAATGACCATGGTTGACACAG 42

RESULT 11

US-09-765-111A-3
; Sequence 3, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PPARGgamma NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2523)
US-09-765-111A-3

Query Match 58.1%; Score 24.4; DB 10; Length 2523;
Best Local Similarity 82.4%; Pred. No. 1.5;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 9 GGCATGCTGGCAGAAATGACCATGGTTGACACAG 42
DB 1075 GCCTCTCTCAGAAATGACCATGGTTGACACAG 1108

RESULT 12

US-09-765-111A-15
; Sequence 15, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PPARGgamma NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 1608
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (91)...(1608)
US-09-765-111A-15

Query Match 57.6%; Score 24.2; DB 10; Length 1608;
Best Local Similarity 78.4%; Pred. No. 1.7;
Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 6 GCAGGATGCTGGCAGAAATGACCATGGTTGACACAG 42
DB 157 GCACCATATCAGAAATGACCATGGTTGACACAG 193

RESULT 13

US-10-109-886-5
; Sequence 5, Application US/10109886
; Patent No. US20020119499A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO. LTD.
; APPLICANT: TANIGUCHI, Tomoyasu
; APPLICANT: MIZUKAMI, Junko
; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND
; TITLE OF INVENTION: ANTAGONIST FOR PPAR
; FILE REFERENCE: TANIGUCHI-6
; CURRENT APPLICATION NUMBER: US/10/109,886
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/514,247
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/JP98/03734
; PRIOR FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: JP231084/1997
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)...(1679)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2003, 23:02:45 ; Search time 1003.67 Seconds
(without alignments)
677.725 Million cell updates/sec

Title: US-09-765-111a-11
Perfect score: 42
Sequence: 1 ccacgcagcagcgtggca.....aatgaccatggtgacacag 42

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Tc number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: em_estba.*
2: em_esthum.*
3: em_estlin.*
4: em_estmu.*
5: em_estov.*
6: em_estpi.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_nam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	71.4	903	9	AL543579
2	30	71.4	1141	14	BM922116
3	24.2	57.6	901	13	BI820841
4	24	57.1	1243	14	BM924484
5	23.6	56.2	426	17	AQ012491
6	23.6	56.2	629	10	BB619013

C	7	23.6	56.2	1658	17	AG184368	AG184368 Pan trogl
C	8	23.4	55.7	318	9	AA161363	AA161363 zq39b10.s
	9	23.2	55.2	594	17	AO614173	AO614173 HS 5127 B
	10	23	54.8	451	13	BI748070	BI748070 rc5e07.y
	11	23	54.8	600	10	BE019930	BE019930 db52d12.y
	12	23	54.8	625	10	BE207571	BE207571 bb66a08.y
C	13	23	54.8	632	13	BG948899	BG948899 QVI-H5003
	14	23	54.8	637	12	BE998291	BE998291 Q01681J83
	15	22.8	54.3	511	17	AZ049319	AZ049319 GSSSRU087
	16	22.8	54.3	590	9	AL826215	AL826215 AL826215
C	17	22.6	53.8	663	13	BI158341	BI158341 602921031
	18	22.6	53.8	689	17	BI110027	BI110027 RPCI-24-3
	19	22.4	53.3	382	9	AA256006	AA256006 zs28c02.r
	20	22.4	53.3	454	13	BI682587	BI682587 463715 WA
	21	22.4	53.3	485	13	BM483972	BM483972 537636 MA
	22	22.4	53.3	485	13	BM484245	BM484245 538020 MA
	23	22.4	53.3	490	12	BF652561	BF652561 276153 NA
	24	22.4	53.3	494	9	AA453686	AA453686 aa18g08.r
C	25	22.4	53.3	500	13	BM288466	BM288466 529740 MA
	26	22.4	53.3	502	12	BF805712	BF805712 QVI-CI017
	27	22.4	53.3	512	14	NA40051	NA40051 yx98a06.r1
	28	22.4	53.3	539	10	AW580428	AW580428 PM2-PT045
	29	22.4	53.3	544	17	AZ286381	AZ286381 RPCI-23-1
	30	22.4	53.3	552	10	AW653703	AW653703 102525 MA
	31	22.4	53.3	566	12	BE740957	BE740957 601595318
	32	22.4	53.3	578	13	BI535582	BI535582 399198 MA
C	33	22.4	53.3	605	9	AI777344	AI777344 EST263752
	34	22.4	53.3	612	9	AI735926	AI735926 tc14d11.x
	35	22.4	53.3	630	9	AI777396	AI777396 EST263804
	36	22.4	53.3	637	12	BF032197	BF032197 601452611
	37	22.4	53.3	661	12	BF212302	BF212302 601813350
C	38	22.4	53.3	683	9	AA496551	AA496551 zv38d01.s
	39	22.4	53.3	692	12	BE747612	BE747612 601577470
	40	22.4	53.3	700	10	BE294158	BE294158 601172919
C	41	22.4	53.3	717	12	BE734591	BE734591 601569924
	42	22.4	53.3	723	12	BE899540	BE899540 601681793
	43	22.4	53.3	760	12	BE742094	BE742094 601575865
	44	22.4	53.3	760	12	BE966552	BE966552 601660231
	45	22.4	53.3	790	12	BF216263	BF216263 601883776

ALIGNMENTS

RESULT 1
AL543579
LOCUS AL543579 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1006Y110 5
DEFINITION prime, mRNA linear EST 16-FEB-2001
ACCESSION AL543579
VERSION AL543579.1 GI:12876058
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 903)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope

FEATURES
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..903

/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="CS0D1006Y110"
/clone_lib="LTI_NFL006.PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Peng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com

BASE COUNT 249 a 228 c 220 g 205 t 1 others

ORIGIN

Query Match 71.4%; Score 30; DB 9; Length 903;
Best Local Similarity 86.8%; Pred. No. 0.75;
Matches 33; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 CGCGCGCTGGCGGAGAAATGACCATGGTTGACACAG 42

|||||
Db 96 CGCGCGCTGGCGGAGAAATGACCATGGTTGACACAG 133

RESULT 2
BM922116 1141 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT 6707162 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5754203
DEFINITION 5', mRNA sequence.

ACCESSION BM922116
VERSION BM922116.1 GI:19372495
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: L1AM12791 row: d column: 12
High quality sequence stop: 105.
Location/Qualifiers

FEATURES

SOURCE

1..1141
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5754203"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"

/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH MGC Library."

BASE COUNT 288 a 291 c 301 g 259 t 2 others

ORIGIN

Query Match 71.4%; Score 30; DB 14; Length 1141;
Best Local Similarity 86.8%; Pred. No. 0.83;
Matches 33; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 CGCGCGCTGGCGGAGAAATGACCATGGTTGACACAG 42

|||||

Db 68 CGCGCGCTGGCGGAGAAATGACCATGGTTGACACAG 105

RESULT 3

BM920841

LOCUS

DEFINITION

BM920841 901 bp mRNA linear EST 04-OCT-2001
603034063F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175351 5', mRNA sequence.

ACCESSION BM920841

VERSION BM920841.1 GI:15932391

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 901)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM11436 row: m column: 16

High quality sequence stop: 876.

Location/Qualifiers

1..901

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5175351"

/clone_lib="NIH_MGC_115"

/lab_host="DH10B"

/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH MGC Library."

BASE COUNT 255 a 214 c 189 g 243 t

ORIGIN

Query Match 57.6%; Score 24.2; DB 13; Length 901;

Best Local Similarity 78.4%; Pred. No. 1.1e+02;

Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 6 CGCGCGCTGGCGGAGAAATGACCATGGTTGACACAG 42

|||||

Db 193 GCAACATATACAGAAATGACCATGGTTGACACAG 229

|||||

RESULT 4

BM924484

LOCUS

DEFINITION

BM924484 1243 bp mRNA linear EST 12-MAR-2002

AGENCOURT 6767565 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5760840

5', mRNA sequence.

ACCESSION BM924484

VERSION BM924484.1 GI:19374863

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1243)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/

TITLE JOURNAL COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12808 row: i column: 01
High quality sequence start: 95
High quality sequence stop: 480.
Location/Qualifiers
1. .1243
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5760840"
/clone_lib="NIH_MGC_116"
/lab_host="DH10B"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon, 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH MGC library."

FEATURES source

BASE COUNT 330 a 376 c 272 g 265 t
ORIGIN
Query Match 57.1%; Score 24; DB 14; Length 1243;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 CAGAAATGACCATGTTGACACAG 42
|||||
Db 292 CAGAAATGACCATGTTGACACAG 315
|||||
RESULT 5
A0012491
LOCUS A0012491.1 GI:3185056
DEFINITION A0012491.1 CIT-HSP Homo sapiens genomic clone 2300P19, DNA
sequence.
VERSION A0012491.1 GI:3185056
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 426)
AUTHORS Adams M.D., Rounsley S.D., Zhao S., Field C.E., Bass S., Linher K.,
Golden K., Berry K., Granger D., Suh E., Wible C., Shizuya H.,
Simon M. and Venter J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL Unpublished (1998)
COMMENT Other GSSs: CIT-HSP-2300P19.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mcadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:

http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES source

Location/Qualifiers
1. .426
/organism="Homo sapiens"
/db_xref="GDB:7154711"
/db_xref="taxon:9606"
/clone="2300P19"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelBAC11; Site 1: HindIII; Site 2:
HindIII"

BASE COUNT 111 a 93 c 121 g 101 t
ORIGIN

Query Match 56.2%; Score 23.6; DB 17; Length 426;
Best Local Similarity 76.3%; Pred. No. 1.4e+02;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 5 CGCAGCGTCGTCGAGAAATGACCATGTTGACACAG 42
|||||
Db 303 CCGGAGCTGGGGCTGAAAGGCCATGATGACACAG 340
|||||

RESULT 6

BB619013/c
LOCUS BB619013.1
DEFINITION BB619013.1 RIKEN full-length enriched, 8 days embryo Mus musculus
cDNA clone 5730466120 5', mRNA sequence.
ACCSSION BB619013.1 GI:15396722
VERSION BB619013.1 GI:15396722
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 629)
AUTHORS Arakawa T., Carninci P., Fukuda S., Furuno M., Hangaki T., Hara A.,
Hiramoto K., Hori F., Ishii J., Ito M., Kawai J., Kono H., Kouda
M., Koya S., Matsuyama T., Miyazaki A., Nomura K., Ono M.,
Okazaki Y., Okido T., Saito R., Sakai C., Sakai K., Sano H., Sasaki
D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Suzuki H.,
Tagami M., Tagawa A., Takahashi F., Takeda Y., Tanaka T., Toya T.,
Muramatsu M. and Hayashizaki Y.
RIKEN Mouse ESTs (Arakawa T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@res.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh
M., Kono H., Okazaki Y., Muramatsu M. and Hayashizaki Y.
Normalisation and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi K., Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E.,
Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura
S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A. and
Hayashizaki Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Kono H., Fukunishi Y., Shibata K., Itoh M., Carninci P., Sugahara
Y. and Hayashizaki Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shinagawa, A., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arakawa, T., Ishii, Y., and Hayashizaki, Y.
 Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Func. Genomics 2 pre, 172-186 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source
 Location/Qualifiers
 1. 629
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="5730466120"
 /clone_lib="RIKEN full-length enriched, 8 days embryo"
 /sex="mixed"
 /dev_stage="8 days embryo"
 /lab_host="DH10B"
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Ret = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGATTATTAATATATCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pluescript KS(+) after bulk excision from Lambda FLIC I. Cloning sites, 5' end: SalI; 3' end: BamHI."
 BASE COUNT 199 a 151 c 139 g 140 t
 ORIGIN

Query Match 56.2%; Score 23.6; DB 10; Length 629;
 Best Local Similarity 76.3%; Pred. No. 1.6e+02;
 Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 TCCGAGCATGGCGAGAAATGACCATGTTGACACA 41
 |||||
 Db 315 TCTCAGAGATGTTCTAGAAATGTCATGATTACACA 278

RESULT 7

AG184368/c
 LOCUS AG184368 1658 bp DNA linear GSS 09-JAN-2002
 DEFINITION Pan troglodytes DNA, clone: RP43-058C03.TJ, genomic survey sequence.

ACCESSION AG184368
 VERSION AG184368.1 GI:16714048
 KEYWORDS GSS.
 SOURCE

ORGANISM
 Pan troglodytes
 Male BAC Library clone:RP43-058C03.TJ.

REFERENCE
 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

AUTHORS
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H., and Sakaki, Y.

TITLE
 BAC end sequences of Library RPCI-43

JOURNAL
 Unpublished
 2 (bases 1 to 1658)

AUTHORS
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Toroki, Y., Watanabe, H., and Sakaki, Y.

TITLE
 Direct Submission

JOURNAL
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

COMMENT

(E-mail: chimpses@gsc.riken.go.jp, URL: <http://hgp.gsc.riken.go.jp/>, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the chimpanzee BAC library RPCI-43. This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

PRIMERS

Sequencing: TJ

LIBRARY

Vector : pBACe3.6
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI.

FEATURES

source
 Location/Qualifiers
 1. 1658
 /organism="Pan troglodytes"
 /db_xref="taxon:9598"
 /clone="RP43-058C03.TJ"
 /sex="male"
 /cell_type="lymphocytes"
 /clone_lib="RPCI-43 Chimpanzee Male BAC Library"

BASE COUNT 352 a 447 c 220 g 545 t 94 others
 ORIGIN

Query Match 56.2%; Score 23.6; DB 17; Length 1658;
 Best Local Similarity 72.5%; Pred. No. 2.4e+02;
 Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CCATCGAGCATGGTGGCAGAAATGACCATGTTGACAC 40
 |||||
 Db 1453 CCATCGAGCATGGTGGTAGAACAGAACATGATGACAC 1414

RESULT 8

AA161363/c
 LOCUS AA161363 318 bp mRNA linear EST 12-MAR-1998
 DEFINITION ZG39D10.s1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone IMAGE:632059 3', mRNA sequence.

ACCESSION AA161363
 VERSION AA161363.1 GI:1735599
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 318)
 Hillier, L., Allen, M., Bowles, L., Duboue, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
 WashU-NCI human EST Project
 Unpublished (1997)

TITLE

JOURNAL

COMMENT

Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Possible reversed clone: polyT not found
 Insert Length: 1847 Std Error: 0.00
 Seq primer: -40M13 fwd. from Amersham
 High quality sequence stop: 257.

FEATURES

source
 Location/Qualifiers
 1. 318
 /organism="Homo sapiens"
 /db_xref="GDB:S182140"
 /db_xref="taxon:9606"
 /clone="IMAGE:632059"
 /clone_lib="Stratagene hNT neuron (#937233)"
 /dev_stage="hNT neurons"
 /lab_host="SOLR (kanamycin resistant)"
 /note="vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT.

Differentiated, post mitotic hNT neurons. Average insert size: 1.5 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGCAG 3' ~3' adaptor sequence: 5' CTCGATGTTTTTTTTTTT 3'

BASE COUNT 59 a 127 c 69 g 61 t 2 others

Query Match 55.7%; Score 23.4; DB 9; Length 318; Best Local Similarity 73.2%; Pred. No. 1.5e+02; Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CATCCAGGCGATGTCGACGAATGACCATGTTGACACAG 42

Db 313 CTGGCAGGCATGAGGACAGCACTTCCAGTGTTCACAG 273

RESULT 9
LOCUS A0614173
DEFINITION HS_5127_B2_P07_T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=703 Col=14 Row=L, DNA sequence.

AC :ION
VERSION A0614173
KEYWORDS GSS.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 594)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380599

COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu

Plate: 703 row: L column: 14
Seq primer: T7
Class: BAC ends
High quality sequence stop: 594.

FEATURES
source Location/Qualifiers

1..594
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate:703 Col=14 Row=L"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at EcoRI sites"

BASE COUNT 192 a 124 c 115 g 152 t 11 others

Query Match 55.2%; Score 23.2; DB 17; Length 594; Best Local Similarity 71.8%; Pred. No. 2.3e+02; Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CCATCGCAGCGATGTCGACGAATGACCATGTTGACA 39

Db 26 CCATTCAGGCGATGCGACCAAAAATACCATGTTNANA 64

RESULT 10
LOCUS B1748070

DEFINITION B1748070.1 GI:15769872
Heterodera glycines.
Heterodera glycines.
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchida; Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
REFERENCE 1 (bases 1 to 451)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Maiza,M., Hillier,L., Kucaba,T., Theising,B., Bowers,V., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

ACCESSION B1748070
VERSION B1748070.1
KEYWORDS GI:15769872
SOURCE EST.
ORGANISM Heterodera glycines.

REFERENCE 1 (bases 1 to 451)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Maiza,M., Hillier,L., Kucaba,T., Theising,B., Bowers,V., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

TITLE The Washington Univ. Nematode EST Project, 1999
JOURNAL Unpublished (1999)
COMMENT Contact: McCarter JP

The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

The library was constructed by Brandi Chiapelli and Dr. James McCarter (bchiapell@watson.wustl.edu & jmcarter@watson.wustl.edu) at Washington University, St. Louis. DNA Sequencing by: Washington University Genome Sequencing Center St. Louis.
High quality sequence stop: 399.
Location/Qualifiers

FEATURES
source

1..451
/organism="Heterodera glycines"
/db_xref="taxon:51029"
/clone_lib="Heterodera glycines J2 pAMP1 v8 Chiapelli McCarter"
/dev_stage="enriched for 2nd stage juveniles"
/lab_host="DH10B"

/note="Vector: pAMP1 (Gibco); Site 1: NotI; Site 2: SalI; The library was constructed by Brandi Chiapelli and Dr. James McCarter at Washington University, St. Louis. The cDNA was made by using Dynabead oligo-dT priming (Dynal). PCR based library using a modified protocol from the SMART PCR cDNA Synthesis Kit from Clontech. Directionally cloned into the UDG sites of pAMP1. Nematodes are the OP50 strain developed by Dr. Charlie Opperman of North Carolina State University (see Genetics, 146:1311-8, 1997). Frozen J2 nematodes were provided by Dr. Rick Davis also of NCSU."

BASE COUNT 127 a 93 c 109 g 122 t

ORIGIN

Query Match 54.8%; Score 23; DB 13; Length 451; Best Local Similarity 74.4%; Pred. No. 2.4e+02; Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 ATCCGAGCGATGTCGACGAATGACCATGTTGACACA 41

Db 251 ATCCGCGAATGCGCGACGAATGACATTTTCGACTCA 289

RESULT 11

BE019930
LOCUS BE019930

600 bp mRNA linear EST 06-JUN-2000

```

DEFINITION   bb62d12.y1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3030071.5'
              similar to gb:X69699 PAIRED BOX PROTEIN PAX-8 (HUMAN); gb:X57487
              M.musculus mRNA Pax8 (MOUSE);, mRNA sequence.
ACCESSION    BE019930
VERSION      BE019930.1 GI:8280021
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 600)
              NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE        Unpublished (1999)
JOURNAL      Contact: Robert Strausberg, Ph.D.
COMMENT      Email: cgapbs@mail.nih.gov
              Tissue Procurement: DCTD/BTP
              CDNA Library Preparation: Ling Hong/Rubin Laboratory
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              image.llnl.gov/image/html/iresources.shtml
              Seq primer: -40RP from Gibco
              High quality sequence stop: 337.
              Location/Qualifiers
FEATURES     source
              1..600
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:3030071"
              /clone_lib="NIH_MGC_9"
              /tissue_type="adenocarcinoma cell line"
              /lab_host="DH10B (phage-resistant)"
              /note="Organ: ovary; Vector: pOTB7; Site:1: XhoI; Site:2:
              EcoRI; cDNA made by oligo-dT priming. Directionally
              cloned into EcoRI/XhoI sites using the following 5'
              adaptor: GGCACGAG(G). Size-selected >500bp for average
              insert size 1.8kb. Library constructed by Ling Hong in
              the laboratory of Gerald M. Rubin (University of
              California, Berkeley) using ZAP-cDNA synthesis kit
              (Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT   160 a 184 c 168 g 113 t
ORIGIN
Query Match   54.8%; Score 23; DB 10; Length 625;
Best Local Similarity 74.4%; Pred. No. 2.8e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CCATCGCAGGCGATGGTGGCAGAAATGACCATGTTGACA 39
Db 132 CCATCGCAGGCGATGGTGGCAGGTAAGAGAGAGCGCGCA 170

RESULT 13
BG948899/c
LOCUS         BG948899          632 bp      mRNA      linear      EST 12-JUN-2001
DEFINITION   QV1-HB0037-300101-591-e07 HB0037 Homo sapiens cDNA, mRNA sequence.
ACCESSION    BG948899
VERSION      BG948899.1 GI:14367070
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 632)
AUTHORS      Dias Neto,E., Garcia Correa,P., Verjovski-Almeida,S., Briones,M.R.,
              Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
              Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
              Brunstein,A., deOliveira,P.S., Bucher,P., Jorgensen,C.V., O'Hare
              ,M.J., Soares,P., Brentani,R.R., Reis,L.F., de Souza,S.J. and
              Simpson,A.J.
              Slotgun sequencing of the human transcriptome with ORF expressed
              sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE     20202663
COMMENT      Contact: Simpson A.J.G.
              Laboratory of Cancer Genetics
              Ludwig Institute for Cancer Research
              Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
              Brazil
              Tel: +55-11-2704922
              Fax: +55-11-2707001
              Email: asimpson@ludwig.org.br
              This sequence was derived from the FAPESP/LICR Human Cancer Genome
              Project. This entry can be seen in the following URL
              (http://www.ludwig.org.br/scripts/gethtml2.pl?l=QV1&t2=QV1-HB0037-
              300101-591-e07&t3=2001-01-30&t4=1)
              Seq primer: puc 18 forward
              High quality sequence stop: 631.
              Location/Qualifiers
FEATURES

```


source

1..632
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HE0037"
 /dev_stage="Adult"
 /note="Organ: bocio,tumor; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 114 a 179 c 209 g 129 t 1 others
 ORIGIN

Query Match 54.8%; Score 23; DB 13; Length 632;
 Best Local Similarity 74.4%; Pred. No. 2.8e+02;
 Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CCATCGCAGCATGTTGGCAGAAATGACCATGTTGACA 39
 |||||
 DL 101 CCATCGCAGCATGTTGGCAGTAAGGAGAGGCGCGCA 63
 |||||

RESULT 14

BE898291 637 bp mRNA linear EST 29-SEP-2000
 LOCUS 601681083F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3951356 5',
 mRNA sequence.
 DEFINITION BE898291
 ACCESSION BE898291.1 GI:10364618
 VERSION EST.
 KEYWORDS human.
 SOURCE Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
 Plate: L10W819 row: e column: 21
 High quality sequence stop: 637.

FEATURES

source

1..637
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3951356"
 /clone_lib="NIH_MGC_9"
 /tissue_type="adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GCGACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 174 a 182 c 170 g 110 t 1 others
 ORIGIN

Query Match 54.8%; Score 23; DB 12; Length 637;
 Best Local Similarity 74.4%; Pred. No. 2.8e+02;
 Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 CCATCGCAGCATGTTGGCAGAAATGACCATGTTGACA 39
 |||||
 Db 113 CCATCGCAGCATGTTGGCAGTAAGGAGAGGCGCGCA 151
 |||||

RESULT 15

AZ049319 511 bp DNA linear GSS 06-MAR-2001
 LOCUS GSSBRU0872 Sheared genomic library Brucella melitensis biovar
 DEFINITION Abortus genomic clone BC09, DNA sequence.
 ACCESSION AZ049319
 VERSION AZ049319.1 GI:7273234
 KEYWORDS GSS.
 SOURCE Brucella melitensis biovar Abortus.

ORGANISM

Brucella melitensis biovar Abortus.
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 Brucellaceae; Brucella.

REFERENCE

AUTHORS Sanchez,D.O., Zandomeni,R.O., Cravero,S., Verdun,R.E., Pierrou,E.,
 Faccio,P., Diaz,G., Lanzavecchia,S., Agüero,F., Frasch,A.C.C.,
 Andersson,S.G.E., Rosetti,O.L., Grau,O. and Ugalde,R.A.,
 Gene discovery through genomic sequencing of Brucella abortus
 Infect. Immun. 69 (2), 865-868 (2001)

TITLE

JOURNAL

MEDLINE

COMMENT

Contact: Zandomeni, R.
 Centro de Investigación en Ciencias Agropecuarias (CICA)
 Instituto Nacional de Tecnología Agropecuaria (INTA) C.C. 25 (1712)
 Castelar, Buenos Aires, Argentina
 Tel: 5411-4621-3316/1683
 Fax: 5411-4481-1316
 Email: zandomeni@inta.gov.ar
 Class: shotgun.

FEATURES

Source

1..511
 /organism="Brucella melitensis biovar Abortus"
 /strain="S-2308"
 /db_xref="taxon:235"
 /clone="Bc09"
 /clone_lib="Sheared genomic library"
 /note="Vector: pBluescript SK(-) (STRATAGENE); Genomic DNA
 was mechanically sheared, blunt ended, and
 size-fractionated by agarose gel electrophoresis.
 Fragments between 1.5-3 Kb were recovered and ligated to
 the EcoRV site of the pBluescript SK (-) vector."
 BASE COUNT 102 a 140 c 142 g 122 t 5 others
 ORIGIN

Query Match 54.3%; Score 22.8; DB 17; Length 511;
 Best Local Similarity 79.4%; Pred. No. 3e+02;
 Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 ATCGCAGCATGTTGGCAGAAATGACCATGTTG 36
 |||||
 Db 54 ATCGGCGCAAGGTGGCTGAACCATGTTGGTG 87
 |||||

Search completed: January 22, 2003, 00:39:39
 Job time : 1006.67 secs